

Towards a world without precedent

The turn of every year is an occasion for reflection, the present more so than most. When last did a year begin with the well-advertised possibility of a damaging war within 15 days? Here are some reflections for the research community.

THE belief that science is international is widely shared, but is it also correct? The guiding principle is that discovery is a common endeavour to which all can contribute and in which all may share. Its formal embodiment is the public literature, to which all have access both as potential contributors and as users, but there is more to say than that. Much of the excitement of the past few years has stemmed from the ways in which previously separate groups of users of this common resource have been able physically to mingle with each other, sometimes choosing to collaborate, sometimes deciding not to do so. The tentative arrangements now being made in Germany for the continued support of eastern German research (see page 3) are a poignant reminder that even much-desired changes may cause hardship; some researchers from old East Germany will be generously supported, others no doubt will lose their jobs. And those circumstances will mark the temper of the scientific community in Germany for many years to come.

So how will research, and intellectual life in general, be shaped by the much greater upheavals under way beyond the German frontiers? There is no way of telling, for the upheavals have still to be completed. Yet some things are already clear. First, the scale of physical migration from one place to another is probably greater than at any previous time in human history. More than a million people have fled the Middle East in the past few months, mostly to Egypt, India, Pakistan and the southwest Pacific, yet the more deliberate continuing flight from the Soviet Union will cumulatively have greater consequences. Second, five new legitimately autonomous states have appeared in Central Europe — and may soon be joined by as many more if last week's elections in Yugoslavia are a guide. Third, it is unthinkable that the Soviet Union can hang together except as a loose association of states at one extreme or as a self-destructive despotism at the other.

Mr Mikhail Gorbachev, the Soviet president, has rightly been given credit (by a Nobel committee, among others) for the reshaping of Central Europe, yet he is now much reviled by his supporters of only a few years ago. His error was to seek to combine the posts of head of state and general-secretary of the Soviet Communist Party, whose influence he rightly sought to diminish. To complain that a man who is manifestly so courageous lacked the courage required by logic at a critical juncture must

seem churlish, but many of the troubles of the past few years have flowed from that mistake. The danger now is that even Gorbachev will rediscover what his predecessors knew, that only the Party's apparatus is used to running the Soviet Union.

Whatever happens, one loser will be Soviet science. Many have left, are leaving or will leave. Many of those remaining are made less productive than they would be elsewhere by the privations of daily life, regularly shown to be pointless by the now-durable connections with the outside world. Yet economic reform, if it now arrived by magic, would make things worse for many people in research, who would find that the establishments to which they are attached could no longer afford them. That is one reason why the research community elsewhere, in its own long-term interests, should now be prepared to offer what help it can. The republic academies need help, as do universities individually. Sadly, finding places for able researchers and promising students has become an urgent need. It will be a great surprise if it is satisfied before the century is out. □

What global treaty?

The chances of negotiating a convention on greenhouse gases are shrinking.

EVERYBODY will agree that if the threat of global warming is real, it must command more urgent attention than any other contemporary environmental issue. This is not to say that adaptation to much higher average temperatures would be technically impossible, but merely that it would be horrendously expensive in money (from the pockets of prosperous populations) and lives — those of the poorest peoples. That, no doubt, is why there has been a substantial measure of agreement among governments on the sequence of events that should be followed in the negotiation of a greenhouse convention (see page 3). Yet optimism about the outcome of this process may be the most serious impediment of success.

The most subtle, but most obdurate, difficulty is that the problem has not been thoroughly talked through. It is customary to say that last year's reports from the Intergovernmental Panel on Climate Change (IPPC) represent what is called a consensus. In the strict sense, this



must be true; presumably panel and subpanel members substantially agree with what their panels and subpanels said. Yet these propositions have not been publicly exposed to the sceptical and even hostile criticism that those who advance unwelcome propositions expect to encounter. On the contrary, there is a sense among the small surviving band of sceptics that their opinions are unwelcome, not least because listening to them would seem to weaken the consensus. In reality, the opposite is the truth. And as long as unrequited scepticism is abroad, the chances are enhanced that governments eager to drag their feet will find expert supporters. IPPC should resolve, for 1991, to engage its doubters in intellectual argument — not avoid them.

The more immediate problem is the gargantuan political difficulty of what the negotiators have in mind. Many governments (the Soviet Union and Iraq, for example), without whom a treaty would be meaningless, are too distracted by other events to pay attention. Then, many suppliers of the international oil market (Saudi Arabia included) have made plain their opposition to a treaty that would restrict carbon dioxide emissions, and would have to be won round. The position of China, the world's largest coal-producer, is unknown. Yet those who ask for a treaty with all the numbers entered have given themselves 18 months in which to find an equitable basis for partitioning the absorptive capacity of the global carbon cycle, for compensating developing countries for development impeded or denied and for deciding how to monitor compliance (and to penalize violations).

But would not all these difficulties melt away if governments had the will to contain global warming? That is what the enthusiasts ask. They are mistaken. Durable treaties do not rest on good intentions, but on devices for exchanging mutually conflicting interests for some mutual benefit. The danger for the next 18 months is that a hasty negotiation will yield a botched treaty, one that the signatories will as readily disregard as keep. The model, in short, should be that followed in the control of gases threatening the integrity of the ozone layer, where a treaty without numbers, but commanding general assent, was followed by more detailed prescriptions. There may, of course, be a price to pay for that. If the predictions are correct, the abatement of a continued accumulation of greenhouse gases will be more difficult, and expensive. The comfort is that the cost of a half-baked treaty that failed would be even greater. □

Education platitudes

If in the 1980s the lamentable state of much public education was acknowledged, will the 1990s see the cure?

THAT many governments are alarmed at the state of the public education they provide is hardly surprising, and for several reasons. First, educated skills (not only technical) are ingredients of economic competitiveness. Second, in a

technically sophisticated world, technical sophistication assists personal fulfilment and is a guarantor of flexibility. Third, many school systems have suffered from neglect or, worse, from too much externally imposed change. (Most explicit expressions of discontent centre on secondary schools; if such discontents could be and were satisfied, higher education would almost certainly be found wanting.)

These propositions are platitudes that are commonly accepted wherever there are worries about public education, but they do not readily provoke remedial recipes. Indeed, when mixed with the proposition that science and technology are essential ingredients in the education of the very young, they lead to recipes whose effects are the opposite of those intended. So much can be learned from the history of the British educational system, which has been (and is still, but less so) encumbered by straitjacket specialization. Students would be invited to declare their lifelong interests before their teens, and would then study more and more of that, and less and less of other learning. One predictable result was to drive away the generalists. Another was to ensure the narrow education of all specialists.

Britain is now bravely launching a national curriculum on its still-astonished secondary schools. When this has been made to function generally, and has settled down, it should yield some of the benefits of the French school curriculum without as much central direction. Yet many of the old errors persist. From being compulsorily narrow, the curriculum has been made compulsorily broad. In particular, everybody will be taught mathematics and science; the hope is that there will be enough specialists to accommodate the needs of the schools. But that is hardly feasible. It would be better for the recruitment of able science students to higher education that science should be left alone than badly taught. (Mathematics is another matter.)

This is where platitudes about the state of public education might have some meaning. Although it is generally agreed that the modern world is more sophisticated than previously, and that almost everybody should benefit from secondary education, it is also generally supposed that this can be done within the same span of a young person's time, and without the erosion of what are called standards. No wonder teachers are up in arms. Governments are asking too much of the educational systems of which they complain. They would serve their own and students' interests better if they set about the imaginative redefinition of a general education and told their secondary schools that that is their prime goal. There would be serious and expensive consequences for higher education, of course, at least where open access is not now the rule. But another relevant platitude is that a good education is not cheap. The financing of public education is certain to be a vivid issue throughout the 1990s. Taxpayers counting on the hypothetical peace dividend to pay for healthcare in old age will have to be encouraged to pay more attention to the needs of the young. □

Global warming makes its mark everywhere

- International treaty planned
- Serious obstacles lie ahead

TAKEN together, 1990 and 1991 may well be best remembered as the years in which the foundations were laid for an international convention to control global warming. An intergovernmental conference at which negotiations will begin is to be held in Washington in February.

At the World Climate Conference at Geneva in November 1990, governments agreed that a convention should be signed at the United Nations Conference on Environment and Development in Brazil in June 1992. But it is not yet clear whether that document will contain binding protocols to limit emissions of carbon dioxide and other greenhouse gases, and to slow deforestation, or whether it will simply be a legal framework on which to build.

Much rests with the United States, which has so far refused to follow other industrialized governments (notably those of Europe) in setting a target to limit carbon dioxide emissions. Throughout 1990, President George Bush stuck to his 'no regrets' strategy allowing for cuts of greenhouse gas emissions only when there are other benefits to be won — energy efficiency, stratospheric ozone protection, smog abatement and the like.

Yet the US position may be influenced by the appearance, perhaps later this month, of the long-awaited Evans report from the US National Academy of Sciences. This will recommend policies the US government should adopt to tackle global warming. The report is expected to blaze a trail beyond no-regrets. With congressional backing, the report could provide the non-partisan weight needed to budge the White House from its opposition to carbon dioxide emission controls.

The climate negotiators will be hoping to build on the environmental achievement of 1990 — the London meeting that tightened the Montreal Protocol's controls on stratospheric ozone-depleting chemicals.

The pledge by the industrialized nations to make available \$240 million by 1993, with more to follow, to help developing countries comply with the protocol, is regarded both as a sign of goodwill and as a precedent for the global warming convention. But the US administration, which inevitably foots a large part of bills like these, is alarmed at the sums of money that might be needed to make a global warming treaty palatable.

The negotiators can at least draw on a solid consensus about the rate of global

warming over the next century predicted by the numerical climate models. The now-completed report of the Intergovernmental Panel on Climate Change (IPCC) suggests that, with no greenhouse gas controls, the average global temperature is likely to rise by 0.3 °C each decade. But reliable regional predictions, and hence a reasonable assessment of the impact of global warming, will require substantially greater computer power and are still a distant prospect.

The economic analysis of strategies to limit and adapt to a changing climate is also largely an uncharted wilderness — such studies as there have been (mostly by

GERMAN REUNIFICATION

Old fears overtaken by doubt

Munich

EUPHORIA, the predominant emotion of 1989 in both halves of the German science community, was replaced in 1990 by a palpable sense of fear. In the West, researchers first feared physical invasion, then being 'infected', from the East. And in the East, people foresaw the end of their institutions and their careers in a brutal wave of rationalizations.

As the year ended, Western researchers felt sheepish but relieved that their worst fears had not come true. Science spending was up, and Western institutes could easily handle the small number of immigrants from eastern Germany.

And as it has turned out, most researchers in eastern Germany have also been spared the axe, at least for a year. That is the effect of the intervention, in July, of the then West German science advisory council (Wissenschaftsrat) and of the West German *Länder* (states), which have jointly resolved on an evaluation of eastern science potential.

Even so, the evaluation exercise has stirred up resentment in the East. Researchers there resent the short time they will be allowed to prove their worth as well as the sometimes-patronizing attitudes of Wissenschaftsrat evaluators.

Once Wissenschaftsrat completes its evaluation in 1991, science in eastern Germany is likely to be given a new lease on life. Towards the end of last year, the granting agency DFG (Deutsche Forschungsgemeinschaft) and the Max Planck Gesellschaft (MPG) took reassuring steps.

Assuming that the newly reelected gov-

advocates of one policy option or another) have produced widely different answers.

Meanwhile, the US administration is also preparing for a long hot future. The Environmental Protection Agency has been quietly expanding its Adaptation Division — an office whose sole purpose is to find ways of living in a greenhouse world. (Canada's new 'Green Plan' calls for the same.) And Department of Energy planners are accelerating their search for alternatives to fossil fuels, while the Department of Agriculture is looking into new crops for what could become a Saharan South and temperate North in the United States.

In some European countries, carbon taxes on fossil fuels are already on the agenda. The European Commission is drawing up plans for a carbon tax throughout its 12 member states. But the British government's recent declaration of opposition to carbon taxes outside the transport sector may cause these plans to be shelved. □

ernment of Chancellor Helmut Kohl continues to give science a high priority despite other pressures on its budget, both organizations promised ample investment in the eastern *Länder*.

DFG has moved boldly and quickly into the East, first (last February) by offering joint grants to eastern and western groups and then, on the day after reunification on 4 October, by allowing eastern Germans to apply directly for DFG money. By the end of the year, more than DM20 million (\$13.3 million) had been granted for 141 cooperative projects. Eastern Germans also applied for DM 15million in direct grants in the last three months alone.

DFG expects nearly DM100 million to be included in its 1991 budget for the eastern *Länder*, nearly 10 per cent of its 1990 budget.

MPG hesitated longer, but announced in November that in 1991 it would support as many as ten eastern research groups. In the process, MPG has also abandoned its policy of independence from university affiliation so as to encourage the resettlement of top groups at eastern universities, which suffered badly under the old regime.

There are no such white knights for the science-oriented industry of former East Germany. Companies such as the optical engineers Carl Zeiss Jena and the electronics manufacturer Robotron may have been lauded, in the past, for the best R&D in the Soviet bloc, but as 1990 ended, they and other companies were about to succumb to old debts and to the desire of most western German industrialists to start afresh in the east. □

Growth continues in adversity

EXCEPT for a few blackspots (Britain, for example), research budgets were dealt with generously last year, despite economic uncertainty. In the United States, a more than usually chaotic budget process produced unexpected largesse. What the future holds is another matter.

Even as late as the summer, most portents in the United States were gloomy. With the congressional appropriations process more bedraggled than ever, and with the prospect that deficit negotiations and partisan infighting would literally paralyse the government by denying it funds, a Congress philosophically supportive of science took to warning of austerity across the board.

By late October, science — especially biomedical science — had magically defied the odds. Support for the National Institutes of Health was increased by more than 10 per cent, or \$850 million, compared with 1990 (which, fiscally, ended in September). The National Aeronautics and Space Administration also did well, with an increase of \$1,646 million, although Congress cut funds for the space station and eliminated all support for a mission to Mars.

But elsewhere, initial optimism turned to disappointment. The National Science Foundation's 13 per cent increase evaporated when, at the last minute, the agency was forced to absorb an unexpected \$40 million in the costs of operating its Antarctic research programme. Physics and materials sciences will suffer as a consequence, with less support, in real dollars, than last year. Worse, the planned five-year doubling of the agency's budget had once again been put off.

Another casualty was the Department of Energy's magnetic fusion programme. The Congress cut \$50 million from the President's request, reducing the programme to 15 per cent below the level of 1990. The result will be no growth — and probably layoffs — in a field once again attractive as a potential long-term palliative to global warming.

Congressional reasoning seems to have been that, if fusion energy is 40 years off, this year's cuts will at worst extend the delay to 41 years. But other fields would suffer more in the short term and thus warrant more immediate priority. Yet the impending loss of fusion researchers is likely to have lasting repercussions.

In Japan, similarly, a perennially straitened government has found the means for a significant increase of direct support for university research. It will be interesting, and important, to learn in due course whether this will be offset, and to what degree, by a reduction of industrial research activity, more than three-quarters of Japan's total.

The research budget has also risen in France by about 8 per cent, substantially more than nominal inflation, which hovers around 3 per cent. But researchers complain that prices of everyday items, such as reagents, have far outstripped inflation. Money at the bench, they say, is even slightly less than ten years ago in real terms.

Although the Mitterrand government has halted and even reversed the previous downward trend in research jobs, anxiety persists that new recruitment is insufficient. Thus the FF10,000 million (\$2,000 million) Centre National de la Recherche Scientifique (CNRS) is worried that young talent is being recruited on too meagre a scale to offset massive retirements later this decade.

Last year, the European Communities (EC) also put together their third five-year framework research and development programme, marked by an expansion into basic research. More than 500 million ECU (about £350 million) has been set aside for the 'human capital and mobility' programme, intended to replace the much smaller current programme of support for basic research, which is due to end in 1992.

But a row broke out when it emerged that European Commission Vice-President Filippo Pandolfi wanted to spend about 80 per cent of the new money on postdoctoral exchanges between EC member states. There was no demand for such a large postdoctoral programme, the member states argued, and more money is now expected to be spent on large laboratories and travel grants to encourage links between established research groups.

In Germany, reunification has created special problems (see page 3), but only Britain seems to be a glaring exception to the rule that industrialized governments seek to spend more, not less, on research.

After two years in which the Department of Education and Science's spending on the research councils increased, the budget for 1991–92, announced in November, threw British science into crisis. Although Education and Science

Secretary Kenneth Clarke described the budget as a "real terms level settlement", outsiders say that inflation will leave the research councils with less to spend next year than this.

The pinch was felt first by the Medical Research Council (MRC), which fared worst in the allocation of the 1990–91 science budget last January. Facing with overspending of £3.5 million before April 1991, MRC froze new appointments and equipment purchases and postponed more than 80 new research grants.

Now, with the announcement of the 1991–92 budget, the Science and Engineering Research Council (SERC) has calculated that it must save £40 million in the coming year and has launched a similar cost-cutting package. Several projects will have to be delayed, including the plan (agreed in principle) to collaborate with the United States and Canada on a scheme to build an 8-metre optical telescope.

The other research councils, more used to financial attrition, have not yet been forced to renege on the consideration of research grant proposals. The Natural Environment Research Council may be protected to some extent by the current popularity of environmental issues and research.

Changes of personnel may matter. The appointment of a biologist, Sir Mark Richmond, as chairman of SERC, has prompted a review of SERC spending on 'big science' and large international collaborations; this is bound to be an important issue in 1991.

Beleaguered British scientists also hope that the replacement of former chemist Margaret Thatcher as Prime Minister by the non-scientist John Major will improve the climate. Later this month, the Conservative MP Sir Ian Lloyd leads a delegation of eminent British scientists, including two members of the well-respected House of Lords Science and Technology Committee, Lords Porter and Flowers, to meet the new man. The possible appointment of a cabinet-rank minister to oversee science and technology will be raised.

At this stage, it is far from clear what the future holds for research spending. In the United States, for example, it was certain that last summer's budget crisis would be repeated in 1991 even before the economic indicators began to point unambiguously towards at least a mild recession, and at worst a full-fledged depression lasting for several years. That can only worsen the federal deficit.

So the outlook for the year ahead is not cheerful. Congress is clearly convinced that science is a critical part of the federal enterprise, and one that should have more money. Faced with layoffs in their home constituencies, however, congressmen's nerves will be tested severely if they are to keep their eyes on this distant prize. □

TOWARDS A NEW YEAR

THIS compilation by regular contributors is intended to remind readers of last year's events and to look forward to the year ahead. Its authors include Peter Aldhous (London), Christopher Anderson (Washington), Steven Dickman (Munich), Diane Gershon (Washington), David Lindley (Washington), Elizabeth Schaefer (San Francisco), Seth Shulman (Boston) and David Swinbanks (Tokyo). □

More problems than products

FOR a nominally fast-growing field, biotechnology had a supremely disappointing year. In the United States, just one new product won government approval (Genentech's gamma interferon, just a week ago). And the year ahead is clouded by fears that the use of genetic engineering to manufacture L-tryptophan in a nutritional supplement may have caused cases of a rare and sometimes fatal blood disorder in the United States (see below).

The US industry during 1990 was preoccupied with a paroxysm of takeovers and mergers, the chief of which was the purchase by Swiss-based Roche of a 60 per cent stake in Genentech, Inc. Other companies were busy forming strategic alliances in Europe and Japan.

Only Genentech and Amgen, Inc. produced meaningful profits in 1990; most US companies have yet to break even. What 1990 showed is that a company's failure to win product approval the first time round can be costly, as Cetus Corporation found when denied approval for interleukin-2 in July. So can protracted patent litigation — witness the interminable patent dispute between Amgen and Genetics Institute, Inc. over rights to erythropoietin.

If dissension in the ranks is the sign of trouble, biotechnology's problems may

get even worse. Last March, Cetus and Genetics Institute noisily quit the Industrial Biotechnology Association (IBA), claiming that the association had taken sides on several key issues. The only significant piece of US legislation, an amendment of the Orphan Drug Act, was given a mixed reception by the industry, then killed by President George Bush.

One genetically-engineered product that might have turned an honest penny was bovine somatotropin (BST), which boosts milk production in cows. But that market dried up when the major milk-producing states in the United States followed the decision of the European Communities and banned the material on economic and health grounds. These decisions were not shaken by the unprecedented step of the US Food and Drug Administration last August in publishing human food safety data while the review of data was still under way. Even the clean bill of health given to the hormone last month by an NIH panel has not exorcised the doubts of consumer lobby groups and animal welfare organizations.

One of the few bright spots in 1990 was the successful first trial of gene therapy to treat a human disease, that of a 4-year-old girl suffering from a rare inherited immune system disorder, adenosine deaminase (ADA) deficiency. A second study, also at NIH, will treat patients with advanced malignant melanoma.

US experience has mostly been mirrored elsewhere, in Europe for example. But there has been some progress with the development of regulations, notably in Germany. Despite the country's conservative stance on embryo research and gene therapy involving the human germline, a law to regulate the use of genetic manipulation in laboratory research and industrial production was passed in June.

Both researchers and pharmaceutical companies are relieved. Hoechst now expects to be able to finish the human insulin pilot plant it had been building near Frankfurt until the courts ordered a halt to the work pending the passage of a law. The technique is, of course, identical with that used elsewhere for almost a decade.

In Switzerland, by contrast, some researchers fear that more restrictive regulations may emerge from discussions now under way in the federal parliament. Proposals that genetic manipulation should be allowed only when it can be shown to be of "immediate medical benefit" have won support in cantonal referenda. The fear is that parliamentary procedures may require these propositions to be reflected in laws due to be drafted in the New Year. Such a move, however improbable, would halt not just biotechnology but molecular biology. □

More monkey business

THE most celebrated icons of the US animal-rights movement — the 'Silver Spring monkeys' — began and ended 1990 (see also page 13). After nearly a decade of legal and administrative fighting, NIH scientists started the year by experimenting on and then killing one of the eight then remaining; in July, three of the rest were similarly killed after experiments, as required by the regulations. By the end of the year, the fate of the remaining four was headed for the Supreme Court.

While the philosophical debate seems no closer to resolution, 1990 may nevertheless be remembered as the year in which the US scientific community mobilized its defence. Both NIH and its sister agency, the Alcohol, Drug Abuse and Mental Health Administration, started new public education programmes to promote the value of animal research. A number of drug companies and medical supply houses made plans for a similar effort of their own.

Animal welfare is another issue on which 1990 proved also to be frustrating. New regulations — the long-awaited amendments to the Animal Welfare Act — were aired for public comment by the Department of Agriculture, which handles animal issues in the United States.

Although final rules for small mammals were eventually approved, the most contentious regulations — those governing dogs, cats and primates — were held up by a lengthy battle inside the administration over the flexibility that should be allowed to research laboratories on the size of cages and exercise requirements. The final regulations, due next month, are likely to be "performance based", meaning that the health of the animals will be the essential criterion, not the details of their housing. Activists consider such regulations unenforceable and are threatening to sue, claiming allies in the Congress.

In Britain, the animal-rights debate took a sinister turn in June when a Ministry of Defence veterinary officer and a University of Bristol neurophysiologist were targeted in car bomb attacks. Neither was seriously hurt. But the violence, and the realization that the more moderate animal rights groups are making inroads with young people, prompted the scientific community to join the debate.

British animal experimenters will have to battle not only for public sympathy, but also for public funds. Universities estimate that upgrading animal houses to comply with 1989 guidelines will cost £75 million, but the government promises only an extra £10 million a year over the next three years. Some facilities may be forced to close. □

Engineering trouble?

THE trouble over L-tryptophan has sent shivers through the biotechnology industry. The material is made by the Japanese chemical company Showa Denko and has been implicated in more than 1,500 cases of eosinophilia-myalgia syndrome (EMS) in the United States, including 27 deaths. If Showa Denko's use of genetic engineering is found to be related to the EMS outbreak, the consequences for the industry could be devastating.

US and Showa Denko researchers have traced the possible cause of the EMS outbreak to a contaminant that appears only during the purification of L-tryptophan; the contamination is not present during earlier stages of manufacture when genetically-engineered bacteria produce the amino acid in a fermentation broth.

But biotechnology is not yet off the hook because it has yet to be shown that the contaminant, di-tryptophan amination acetaldehyde (DTAA), causes EMS. Nor has it been firmly established that the appearance of DTAA during purification is unrelated to the use of genetically-engineered bacteria earlier in the manufacturing process. Animal experiments now under way in the United States and Japan may help resolve the issue. □

Battling for and against rising numbers

LAST year may have been that in which the world (except Japan) discovered that there is something amiss with the state of public education, but only in France did the victims — *lycée* students — take to the streets. Elsewhere, committees were convened instead.

The Socialist government of Prime Minister Michel Rocard was clearly shaken by November's demonstrations — the first upheaval in an otherwise tranquil administration. They have been taken as a sign that education should have become a "national priority" ten years ago. Yet this government and its education minister, Lionel Jospin, have set out to make real changes in the educational system. Intended reforms, it seems, are insufficient.

The plans for university reform are similarly striking. Jospin's own specialist adviser, Claude Allègre, is now putting the finishing touches to a long-term plan. A new research and doctoral studies directorate (DRED) has been set up, and aims to double the number of PhDs awarded each year.

Meanwhile, an Institut Universitaire de France is to be created as a sort of virtual Institute for Advanced Studies to help 15 of the best senior professors and 25 younger professors to carry out research.

Will French universities one day become major teaching and research institutions as in the United States and Britain, rather than runners-up to the *grandes écoles* for undergraduates and to the *grandes organismes* (such as CNRS) for research? That seems to be the goal.

As a stopgap measure, Jospin is building new universities (some are already well under way) to modernize and expand the crumbling tower blocks put up hurriedly in the previous expansion of the 1970s. But the questions remain whether building can keep pace with rising student

intake and, with some 1,500 lectureships vacant, where the manpower will come from.

In the United States, the gathering clouds of an economic storm darkened 1990 for the research universities in a potentially painful way. The Congress, struggling with the federal deficit, sent the federal science agencies the message that indirect costs — the supplement (as much as 80 per cent) of federal research grants that goes to pay overhead costs to research institutes — have grown, are too high and must be reduced.

Congress set a 14 per cent cap on the indirect costs of research grants in agriculture and put pressure on the National Institutes of Health (NIH) to promise action in its own field. Representative John Dingell (Democrat, Michigan), the dogged congressional watchdog, closed 1990 with a withering examination of accounting methods at Stanford University. He is planning early this year a public investigation of indirect costs as a whole.

The critics of the system have found unexpected allies among university researchers, some of whom tend to blame their institutions for the increasing difficulty of winning individual research grants. They note that while the research portion of NIH grants grew by less than 10 per cent over the past decade, the associated indirect costs grew by nearly 25 per cent.

Meanwhile, the general economic downturn in the United States threatens to push universities even closer to ruin. Tuition rates are rising out of students' reach and endowments are declining. There is rough water ahead.

The reunification of Germany could cause a nightmare of a different kind in 1991 if, as can hardly be avoided, the increased number of students brings more crowding and even longer courses.

With its eyes on the east, the German government has made no progress in relieving the pressure on chronically overcrowded universities by hiring more faculty or restricting access. The situation will be exacerbated when thousands of young people in the the eastern *Länder* begin to matriculate at the better equipped universities in the west. Matters will be made even worse if, as expected, investment in the west is slowed so as to provide more money for modern lecture halls and laboratory benches in the east. It is a fine calculation whether these developments — and the lack of cash — will serve to keep the easterners close to home.

British universities, as always, are in the most precarious condition. 1990 brought yet more upheaval and confusion, principally as a result of the collapse of the attempt by the Universities Funding Council (UFC) to introduce a system of competitive bidding for funds to support undergraduate teaching. The idea was that universities wishing to expand would take advantage of supposed 'economies of scale' and teach more students at a lower cost per student. The outcome would have determined student numbers and corresponding incomes at different universities for the four years 1991–95.

In the event, the universities refused to play ball, arguing that the UFC's "guide prices", meant as maximum costs in various subject areas, were too low. Bids to teach extra students did arrive, but the vast majority were at the guide prices. In October, the UFC was forced to admit defeat, suspending the bidding system.

What happens next is unclear. (The change of prime ministers is probably not irrelevant.) The UFC says it will allocate student places and teaching funds for 1991–92 in February, and give provisional figures for the succeeding three years in March. But the council has too little money for the planned expansion at the present cost per student. Many universities fear that they will be forced to charge 'top up' fees to students to maintain the quality of university education.

The UFC itself is believed to be divided over how to proceed, although whoever replaces the retiring chief executive, Sir Peter Swinnerton Dyer, in the Spring faces an unenviable task in restoring the UFC's battered public image.

It could well be that the central government will seek to finesse UFC's difficulty by broadening the argument, perhaps by reopening the question of whether the distinction between universities and other institutions of higher education, notably the lower-cost polytechnics, should be drawn as rigidly as at present. British universities could yet find that they have won a famous battle, but not the long war of attrition in which they have been engaged for the past two decades. □

'TOP 10' US UNIVERSITIES: BIOLOGICAL AND PHYSICAL SCIENCES (CITATIONS PER PAPER)

BIOLOGICAL SCIENCES		PHYSICAL SCIENCES	
1. Rockefeller University	7.96	1. Univ. Calif., Santa Cruz	4.56
2. Caltech	7.71	2. Harvard University	4.21
3. MIT	7.04	3. Princeton University	3.76
4. Stanford University	6.19	4. University of Chicago	3.74
5. Princeton University	6.07	5. Univ. Calif., Santa Barbara	3.57
6. Univ. Calif., Berkeley	5.96	6. Yale University	3.35
7. Harvard University	5.61	7. Boston University	3.30
8. Univ. Calif., San Francisco	5.05	8. Caltech	3.28
9. Univ. Calif., San Diego	4.59	9. Stanford University	3.27
10. University of Oregon	4.58	10. University of Houston	3.09

These are the top universities in the United States by the yardstick of 'research quality' as determined by the average number of citations per paper in the biological sciences (biology and clinical medicine) and physical sciences (physics, chemistry, earth sciences, engineering, mathematics, and applied sciences). The figures, taken from journals indexed by the Philadelphia-based Institute for Scientific Information between 1987 and 1990, will appear later this year in *Science Watch*. (Data: ISI's Science Indicators Database, 1987–90.)

Benefits not appreciated

DESPITE the promise of nuclear energy in the avoidance of global warming, 1990 did little to resolve the continuing argument about the prudence of building nuclear plants. And the chances are that 1991 will leave most matters unresolved.

Yet the debate has at least begun. In Sweden, for example, the earlier decision that nuclear power would be phased out when existing nuclear plants reach the ends of their lives has been questioned (mostly by industrialists concerned to know where alternative energy sources would be found). The argument is likely to become lively before it is settled.

In the United States, a similar debate also began in earnest, mostly centred about the potential of new reactor designs. It might have been more constructive if it had not been consistently overshadowed by reports of environmental contamination at US nuclear weapons production facilities.

There were, for example, new fears of chemical explosions in the underground tanks holding high-level nuclear waste at the Hanford reservation. There was little progress during 1990 at the Yucca Mountain site earmarked for the disposal of high-level commercial waste, or at many of the sites designated for low-level waste which are required by law to be in operation by 1993.

Some vestigial controversies have nevertheless been resolved. The Environmental Protection Agency in 1990 finally approved a testing period for the salt-bed repository in New Mexico intended for military transuranic waste. And the long-contested Seabrook nuclear power plant in New Hampshire finally won its operating license.

Other long-standing controversies have been revived in the United States. During 1990, interest in the health threat posed by low levels of radiation was confused by conflicting results from research. A study completed this year by the National Cancer Institute found no increased incidence of cancer among neighbours of selected nuclear reactors across the United States, but there is epidemiological evidence of elevated leukaemia rates among those living near the Pilgrim power plant in Massachusetts.

In Japan, despite public opposition to nuclear power since the Chernobyl accident, the Japanese government seems determined to carry on with its ambitious plans to expand nuclear power. But the government will face a political test of its policy in February, when an anti-nuclear power candidate will be among those running for the post of governor in Aomori Prefecture on the northern tip of the main island of Japan, where a huge complex for reprocessing and enriching nuclear fuel

and for storing low-level and high-level waste is being built.

The complex has become a focus of a nationwide campaign by anti-nuclear activists. Although many living in the immediate vicinity of the complex have been mollified by the government's generous grants for new local facilities, many farmers and fishermen nearby are still opposed to the plant, fearing that sales of their products will be reduced.

Meanwhile, Europe's top nuclear nation continues to test missiles and their warheads in the South Pacific (and caesium-134 has been found in the surrounding sea). At home, although the nuclear industry gives France a good record on carbon dioxide emissions, the security of its reactor plants and waste treatment installations has been questioned by a string of minor, but potentially serious, oversights.

In Britain, the nuclear industry must be glad to see the back of 1990. The government's ill-conceived and subsequently aborted plan to sell British nuclear power stations to private investors was attacked in a damning report from the House of Commons Energy Committee, which did much to uncover the mystery of how the costs of nuclear-generated electricity spiralled in the run-up to privatization. The new government-owned nuclear generating companies face another year of stagnation, with Britain's nuclear programme on hold, pending a review in 1994.

The UK Atomic Energy Authority, now trading as AEA Technology, lost money, closed research reactors and shed staff in 1990, and is placing its hopes on diversification away from the nuclear industry.

Perhaps the biggest blow came in February, when a controversial study by a team led by Martin Gardner, from the University of Southampton, linked the excess of childhood leukaemia cases around British Nuclear Fuels Limited's (BNFL)'s Sellafield waste reprocessing plant to their fathers' exposure to radiation at the plant. BNFL will take some comfort from the subsequent failure to find excess cancer cases around nuclear sites in the United States and France.

The deepest confusion on nuclear power is in Eastern Europe. What looked at first like a wide-open market for Western nuclear plant constructors is proving to be a tough sell.

Germany, which abandoned domestic reprocessing of nuclear fuel in 1989, made no move last year to build new nuclear plants in its eastern regions. Instead, the five Soviet-built plants, which used to provide more than 10 per cent of East Germany's electricity, have been shut and

replaced in stopgap fashion with gas-fired and oil-fired plants from the west. Despite Germany's intention to reduce both carbon dioxide emissions and dependence on foreign oil, the plants are likely to remain in place. Building new nuclear plants is just too unpopular.

Hungary and Czechoslovakia are in a similar position. They have so far resisted proposals from the West that they should provide sites for modern nuclear plants that would be paid for by deliveries of electricity to the West.

The decision in mid-1990 to close the Soviet-designed plants at Greifswald, Germany, followed a Western safety study that revealed fundamental design and maintenance flaws. The same flaws would surely be found at identical plants in Czechoslovakia, Bulgaria and the Soviet Union, none of which can afford to shut them down. Indeed, the plant at Kozloduy accounts for more than 40 per cent of Bulgaria's electrical generating capacity. A negative outcome of the safety study now being conducted by the International Atomic Energy Agency and due to be completed this year could throw Bulgaria into an energy crisis.

In Germany proper, the defeat of the Green Party in the December Bundestag elections by no means signals the end of the antinuclear movement. The first battle of 1991 will come over the perennial nuclear waste problem; the dump sites selected by both East and West Germany are considered inappropriate by nuclear opponents and are viewed just as critically by the government of Lower Saxony, in which the Greens are a junior partner.

There is also a potential nuclear scandal brewing over the former Soviet uranium mining areas near Aue in Saxony and Gera in Thuringia. How badly was the local population contaminated? No one knows, but there are plenty of people anxious to find out. □

ANTARCTICA

A deep depression

LAST year was a disappointment for many Antarctic researchers, who know that 1991 may be no better. The 39 Antarctic Treaty nations, at a three-week meeting in Chile that finished last month, made a start on a new protocol to protect the Antarctic environment.

But much remains to be done, starting at a meeting in Madrid in April, where the divisive issue of future minerals exploitation in Antarctica will be raised.

Japan, Britain and the US would like to keep open the option for future mining, but other nations, led by Australia and France, want a permanent ban. Some Antarctic scientists are concerned that a new regime for environmental protection in Antarctica may restrict research. □

The cost of ambition

NINETEEN NINETY was a better year for science in space than public perceptions suggest. In the United States, the new taste for bashing the National Aeronautics and Space Administration (NASA), begun after the Challenger disaster, was fed last year by tales of the Hubble Space Telescope's flawed mirror, the glitches that interrupted Magellan's video transmissions from Venus and the burned-out computers and plumbing problems that hampered Astro-1's mission. But in the end, a great quantity of data was recovered, even from the damaged Hubble.

Galileo and Ulysses were sent off with hardly a hitch on their interplanetary voyages (Galileo to Jupiter via Venus, Ulysses to the Sun via Jupiter). The flaw in Hubble's mirror, eventually traced to a misassembled test device used in its construction, seemed to indicate more than anything vast hubris on the part of NASA and its contractor: so sure was everyone that the construction method was perfect that evidence of error was simply ignored. But that was NASA ten years ago, when Hubble's mirror was built.

Now, things seem to be running more smoothly (with ten successful shuttle launches in a row). To what end? The space station, a politically inspired ambition to which NASA dutifully clings, is losing support because of both its expense and its apparently infinitely redefinable purpose. The European Space Agency (ESA) is more or less happily on board, after earlier ructions, but scientific support for the venture is negligible on two continents.

So is NASA in better shape? Last month's Augustine report chides NASA for continuing bureaucratic inflexibility, lack of imagination and lack of purpose. NASA, apparently, would interpret the report — mostly a bluntly worded criticism — as an endorsement of its present practice. Congress is free with criticisms of NASA, some of them tangible. In last year's budget-haggle, it cut back the space station and put the Mars-Moon mission in hibernation. The agency now sails ahead on the winds of lingering White House support. How long that can last seems open to question.

European space science, meanwhile, enters 1991 in a buoyant mood. After a valiant salvage effort, the Hipparcos satellite should now yield most of the astrometric data originally hoped for, even though it failed to reach a geostationary orbit after its launch in 1989. And the much-delayed Ulysses mission to view the poles of the Sun was finally launched by the shuttle Discovery in October.

Independently of ESA, German and British instruments on the Rosat satellite, launched by the United States during the

Summer, are revolutionizing X-ray and extreme ultraviolet (XUV) astronomy. By February, about 1,500 XUV sources will have been identified, compared with the handful previously known.

Meanwhile (last month), the ESA council resolved the long-running argument over the Horizon 2000 space science budget. Although no Horizon 2000 launches are planned for 1991, two important remote sensing satellites should reach orbit. Meteosat, a weather monitoring satellite, is due for launch in February, and will be followed in April by ERS-1, which has an important role in gathering data bearing on global change. Both are part of the grandly-named 'Mission to Planet Earth'.

Towards the end of 1991, Ulf Merbold, a German astronaut, will conduct a series of microgravity experiments in the Spacelab facility carried by a US space shuttle. The International Microgravity Laboratory mission involves ESA, NASA and the Japanese, Canadian, German and French space agencies.

One uncertainty ESA still has to resolve is the future of its main contribution to the Hubble Telescope: the Faint Object Camera (FOC), whose performance has been impaired by the spherical aberration of the primary mirror. One option being considered is a plan called COSTAR, not yet accurately costed, to replace one of four instruments at the base of the telescope by an arm inserting corrective optics in front of the other three. A decision is expected this year.

In France, the dynamo of European space technology in recent years, what may be the first signs of disenchantment are emerging. France has been backing manned space — through the Hermes spaceplane and Ariane 5 launcher — mostly because it does not want its (and Europe's) aerospace and electronics industry to be left behind.

But there are flies in the ointment. Hermes is costing much more than planned and, in March 1991, the ministerial council meeting of ESA will have to decide

whether Hermes and its launcher, Ariane 5, can proceed to the construction phase.

French scientists are not unanimous about the value of manned space. The Académie des Sciences has twice rejected the idea.

That same ministerial council, or its successor in June, will also have to make a final decision on the future of Columbus, the proposed European module of the US space station. Although transatlantic relations on space projects have recently improved, the rules allow any one of ESA's 13 members to withdraw if costs have increased by more than 20 per cent above the original estimates. Without exceptionally creative accounting, that licence is likely to be on offer.

Japan's National Space Development Agency (NASDA) is plagued with problems of a different kind in its development of its next-generation H-II rocket: the first-stage liquid fuel engine keeps bursting into flames during test runs, each time for a different reason.

The first launch had already been postponed a year (until 1993) after a fire in 1989. And there were two other major fires in 1990. NASDA officials say they are still aiming for a first launch in 1993, but the target seems steadily to recede.

In contrast, Japan's small academic organization, the Institute of Space and Astronautical Science (ISAS), had a very successful year, launching Japan's first spacecraft to the moon. The tiny satellite Hiten (named for a Buddhist angel) has swept past the Moon several times since January in a complex series of Earth-Moon orbits designed to test the "swing-by" technique. In March, Hiten will make a high-speed pass through the Earth's upper atmosphere to test "air-braking" techniques. It may then be parked at one of the Lagrange points where the gravitational forces of the Earth and Sun cancel.

In August or September, ISAS plans to launch the satellite Solar-A into low Earth orbit to observe solar flares near the solar maximum using a Japanese hard X-ray telescope and US soft X-ray telescope. It all sounds like great good fun. Maybe ISAS knows something that NASA (not to mention NASDA) does not. □

Super Collider collision ahead

ONE controversy likely to be prolonged into 1991 concerns the contribution to the Superconducting Super Collider (SSC) requested of Japan in June by the US Department of Energy (DoE).

In Tokyo, government officials are still brooding on a response. One official of the Ministry of Education, Science and Culture (MESC) says he and his colleagues wanted to "run away" when the DoE delegation arrived, knowing that MESC's shoestring budget for science simply cannot cover the costs of such a project. Another, from the Science and Technology Agency (STA),

believes that Japan will not contribute to the construction costs, although some physicists are keen to participate in experiments.

Japanese government officials give the impression of lying quiet in the hope the SSC will just blow away. On the other hand, prominent Japanese physicists are miffed that they have not been asked for their opinions of SSC.

The whole business has drawn attention to Japan's lack of a central body to deal with major issues of international science and technology policy. □

Eighteen ninety one and all that

J. L. Heilbron and W. F. Bynum

A celebration of scientific anniversaries featuring dancing frogs' legs, Java Man, the birth of electromagnetism, French metrication, the classification of screws, a treatise on the astrolabe, and much, much more.

WE list below some attractive opportunities for the celebration of anniversaries of significance in science and technology. Our pick of the year is the 200th anniversary of the publication of Luigi Galvani's *De viribus electricitatis in motu musculari*, in which the dances of disembodied frogs' legs were first made known to a wide audience. Galvani's conviction that the kicking arose from an electricity peculiar to animal bodies came under attack by Alessandro Volta; the dispute lasted a decade, until Volta succeeded in duplicating the power plant of the frogs' legs by a 'pile', or battery, containing no animal matter. The battery gave birth to electromagnetism, and thence to motors, dynamos, telegraphs — in short, to modern Western civilization.

Runners-up for the anniversary of the year are the death of Robert Boyle (1691, a 3-centenary or 3.0¢ celebration) and the foundation of the Society of Jesus (1541, 4.5¢) by Ignatius Loyola (born 1491, 5.0¢). Boyle, probably known best to readers of this journal as a chemist, promoted a vision of a 'Christian virtuoso', that is, an observant Protestant natural philosopher; the Jesuits likewise developed a mix of science and religion, in their capacity as schoolmasters to Catholic Europe. A great many, perhaps the majority, of French and Italian mathematicians and physicists of the 17th and 18th centuries were educated by the Jesuits. Both Galvani and Volta passed through their hands, and Volta almost joined them. Even in the land of Christian virtuosos the Order's teaching had admirers. "As for what pertains to pedagogy," wrote Sir Francis Bacon, "it will be most briefly said: consult the schools of the Jesuits . . . If only they were ours."

Once every century in our millennium the reckoning *anno domini* results in a palindrome. This is the year for our century — 1991. In recognition whereof, we have noticed palindromes back to 1221. Our next offering of palindromic anniversaries will occur 11 years from now, in 2002; thereafter, no one will be able to enjoy two sets of these remarkable observances for a thousand years.

We begin with events of 1891 and move backwards at 50-year intervals. We end with eligible dates from our century.

1891 (1.0 centenary). It was a year of links. Eugene Dubois found bones in Java that he named *Pithecanthropus erectus* and supposed to belong to a creature intermediate between apes and man; Richard Semon, a professor from Jena, examined the habitats and sex life of *Ceratodus forsteri*, an Australian lungfish, which he declared to stand or swim

a 20-pound, 1-horsepower steam engine. Hermann Ganswindt had another idea. He would propel an airship by firing a cannon out its rear. For the vindication of his approach, see below under 1941. Continuing to drop our gaze, we record the invention, by James Dewar and Frederick August Abel, of cordite, the high explosive that the British would favour during the First World War.

The *Princesse Alice II*, the first ship designed and equipped for oceanographic work, was launched by Albert I of Monaco in 1891. The ship did its business so well that the principality soon required a maritime museum and laboratory. It went up in 1899, not, as erroneously reported in this column two years ago, in 1889.

The 'electron' — the name at least, if not the thing — is a century old. The word was coined by the Irish physicist George Johnstone Stoney (died 1911, 0.8¢), following up suggestions made by Hermann von Helmholtz in his 1881 (1.0 palindrome) Faraday (born 1791, 2.0¢) Lecture that electricity consists of discrete charges ('atoms of electricity') and that chemical forces are, at bottom, electrical. Wilhelm Eduard Weber, who gave more than a word to electronics, died in 1891. He had 'found' a law of force between two elements of electric current that

Galvani's experiment — making frogs' legs dance.

between fish and amphibians; and William Matthew Flinders Petrie, digging around in Mycenae, established correspondence between the chronologies of the kingdoms of Egypt and the civilizations of the Aegean.

This year a century ago Max Wolf of Heidelberg turned his camera to the asteroids. He found the 323rd, the first detected photographically; his technique proved fertile, and netted him 500 more. The easy comparison of object with background made possible by the camera had yielded, a decade earlier (1881, first palindrome), the first photographic discovery of a comet. The discoverer, Edward Emerson Barnard, received \$200 for his trouble, from a philanthropist partial to comets.

Closer to the ground, Samuel Pierpont Langley announced, prematurely, that "mechanical flight is now possible with engines we now possess". He had in mind

included Coulomb's inverse-square electrostatics and the electromagnetic effects discovered by Ørsted and Faraday. H. A. Lorent's electron theory may be considered a blend of Weber's approach with Maxwell's field theory.

Medical scientists were still coming to terms with the state of bacteriological discoveries of the previous decade, and trying to move from the diagnosis to the positive treatment of infectious diseases. Following Emil von Behring's announcement of antitoxin therapy, mentioned last year, Paul Ehrlich showed that plant toxins could also be antigenic and reported excellent results treating patients suffering from malaria with methylene blue, a vital stain and mild antiseptic. Although the line of research ultimately bore fruit, it did not immediately oust the old methods of the Jesuits, whose bark (chinchona, whence quinine) continued as the treatment of choice. Sven Hedin made

Mary Evans

diagnosing anaemia easier with the announcement of his 'haematocrit', and Walter Wynter and Heinrich Quincke independently introduced lumbar punctures for both therapy and diagnosis.

Our anniversarial university this year is Stanford, which opened its doors a century ago. Its initial endowment was provided by the railway magnate and politician Leland Stanford, in memory of his

generators, the distribution net, and the lamps and other loads, the station demonstrated the feasibility and the profitability of widely available electrical power. The rest has been history.

1841 (1.5 centenary). We begin small. 150 years ago, Joseph Whitworth introduced his system of classifying screws. The resultant standardization of pitches assisted the perfection of interchangeable parts,

gases and liquids to very severe pressures, and Alfred Cornu, of the Cornu spiral, who studied the gentler subject of optics.

1791 (2.0 centenary). We have already noticed, as the anniversary of the year, the announcement by Luigi Galvani of his ways of stimulating decapitated frogs. A popular new science resulted. Beyond the Alps, decapitation began to enjoy a vogue among higher creatures when the Constituent Assembly, by vote of 6 October 1791, declared that thenceforth all executions in France would be accomplished by beheading. Thus blossomed the proposal made in 1789 by Joseph Ignace Guillotin, who, in the democratic spirit of the times, argued the unfairness that allowed only nobles to suffer capital punishment literally. There is nothing new under the sun. Nobles have enjoyed the services of crude guillotines since the 16th century.

Another useful device from Revolutionary France is the meter stick (see *Nature* 348, 105–106, 1990). Two months before it adopted the guillotine, the Assembly put at the disposal of the Paris Academy of Sciences a large sum — about twice its pre-Revolutionary salary budget — to measure a meridian of the Earth through Paris from Dunkirk to Perpignan. From this measurement, the length of an average degree of that meridional arc would be calculated; and one ten-millionth of 90 times the average degree would serve as the basis of a new, universal and natural system of weights and measures. The system had its difficulties, in execution and implementation; the people, to whose interest rhetoric ascribed its invention, declined to accept it; and the government failed to impose it until 1840 (1.5¢, to better than 1%).

Elsewhere, William Gregor (born 1761, 2.3¢) discovered titanium; Robert Gray, the Columbia River; Pierre Prévost, his theory of heat exchanges; Nicolas Leblanc, his process for manufacturing soda; and Jeremias Benjamin Richter, his law of equivalent proportions.

1791 also saw the births of Charles Babbage and Michael Faraday, both almost runners-up for anniversarians of the year; of Felix Savart, previously mentioned (died 1841, 1.5¢); of Alexis Thérèse Pétit, of the law of Dulong and Pétit; of Johann Franz Encke, of the comet; of Robert Knox, of Burke and Hare notoriety; and of Ottaviano Fabrizio Mossotti, an able exponent of the laplacian approach to electrical theory against the field-theoretic concepts of Faraday.

1771 (2.0 palindrome). Our second palindromic year was distinguished by the first full publication of the first edition of the *Encyclopaedia Britannica*. Much of it was compiled, in the time-honoured way, with paste-pot and scissors, by the man who also printed it, William Smellie. His dictum was that "utility ought to be the principal intention of every publication",

Australian lung-fish — living fossil somewhere between fish and amphibian?

son Leland, Jr. For one of its accomplishments, see below, at 1941.

Gabriel Lippmann (died 1921, 0.7¢) invented colour photography through the interference of standing light waves, for which he received the Nobel prize for physics in 1908; Emil Fischer established the structure of sugars, for which he received the Nobel prize for chemistry in 1902; Edward Goodrich Acheson (died 1931, 0.6¢) invented carborundum, for which he received no Nobel prize. The year 1891 was not good for mathematicians. There died the famous Sonay Kovalesky, a genius at differential equations, also a novelist and a strong voice for equality for women; Leopold Kronecker, whose fortes were whole numbers and elliptical functions; and François-Édouard-Anatole Lucas, who found the largest Mersenne prime ($2^{127}-1$) ever uncovered without electronics, claimed to have advanced the proof of Fermat's last theorem, and wrote a four-volume treatise on mathematical recreations. Carl Wilhelm von Naegeli's mathematics, altogether more prosaic, consisted in representing the apical cell divisions of vascular plants in terms of equations.

1881 (1.0 palindrome). It is gratifying to be able to suggest a world-historical scientific-technological success for our first palindromic celebration. The event was the opening of the Pearl Street Station in Lower Manhattan, the first plant for the centralized distribution of electrical power. Perhaps the greatest of all the inventions of Thomas A. Edison (died 1931, 0.6¢), who designed the motors and

at least in English measure; and the exactness of work on which the scheme rested enabled Whitworth to outdistance all other makers of machine tools at the Great Exhibition of 1851 (1.4¢). He also made important contributions to that leading sector of precise manufacture, armaments. The fortune he gained from screws and guns went mainly to educational and charitable purposes around Manchester. There, back in 1841, his fellow adopted Mancunian, James Braid, was likewise engaged in work equivocally helpful to humanity: the transformation of mesmerism into hypnotism.

Still in the productive English provinces, an association of railroad engineers introduced the lighting convention still used: red for stop, green for slow ahead, white for all clear. Far from the Midlands, James Clarke Ross, in Her Majesty's ships *Erebus* and *Terror*, reached 78°11' South latitude, and determined the location of the South magnetic pole at 75°S, 154°E, about where Carl Friedrich Gauss had reckoned it without leaving Göttingen a few years before. Further to geophysics, the American meteorologist James Pollard Espy published his cantankerous *Philosophy of storms*, which contains his ideas about the thermodynamics of clouds.

The year 1841 brought the deaths of Nicolas Clément and Felix Savart (born 1791, 2.0¢), two outriders of the so-called laplacian school of physics, best known, respectively, for measurements of specific heats and for investigations on electromagnetism; and the births of Emile Amagat, who made his career subjecting

achieved in this case by summing knowledge into treatises, for example, under 'anatomy', rather than parcelling it out in bits arranged alphabetically. "Where is the man who can learn the principles of any science from a Dictionary?" In the event, the *Britannica* had to compromise. Material that lent itself to systematic development was presented in treatises, the rest in dictionary-like snippets.

The otherwise irreproachable first *Encyclopaedia Britannica* has no entry for 'centenary' or 'centennial'. Johnson's dictionary does not have the latter and does not admit the former in the sense of 'anniversary'. The *Oxford English Dictionary* gives a quotation from 1788, which mentions an observance of the 100th anniversary of the Glorious Revolution as its earliest example of 'centenary' in our sense; 'centennial' was so used first in 1797. We shall elaborate the meaning of these facts on another occasion: here we remark only that, since the fundamental concept of our art is scarcely two centuries old, we may be excused if it has not yet attained the perfection of geometry.

1741 (2.5 centenary). 1741 saw the publication of Johann Peter Süssmilch's *Die göttliche Ordnung in den Veränderungen des menschlichen Geschlechts aus der Geburt, dem Tode und der Fortpflanzung der selben erwiesen*. This revelation of God's dispensation in statistics of births and deaths became a foundation text in *Cameralwissenschaft*, a technical political science that formed the bureaucrats of the *Aufklärung*.

The Royal observatory in Greenwich experienced the occultation of Edmund Halley, who retired in 1741, and the arrival of his successor, James Bradley, the discoverer of the aberration of starlight and the nutation of the Earth. While Bradley was settling in at Greenwich, Vitus Jonassen Bering, a Danish explorer in Russian service, took ship across the Straits now named after him and landed in Alaska. Returning West, he sailed past (and so discovered) the Aleutian islands and ran into Bering Island (Ostrov Beringa), which he mistook for Kamchatka. He died there on 8 December.

1691 (3.0 centenary). There perished simultaneously, in the body of the Honourable Robert Boyle, a Christian virtuoso and a sceptical chemist. The sceptic in him ridiculed the received theories of the peripatetics, the alchemists and the paracelsians (see at 1541); the chemist tried all the old and new recipes, and advanced tentative explanations of their workings (see at 1661); the virtuoso ornamented that great collection of virtuosi, the Royal Society of London; and the Christian devoted a portion of his large fortune to proselytizing for his church at home and abroad. Boyle's combination of piety and morality in the personal sphere, and responsibility and inquisitiveness in

matters of science, had a parallel in the naturalist John Ray, whose *The wisdom of God manifested in the works of creation* appeared in the year of Boyle's death.

Other casualties of 1691 were Adrien Auzout, a French astronomer, who helped perfect the micrometer eyepiece (see 1641); and Richard Lower, who

tables', as corrected by himself.

The stars danced at several scientific births that year. Appropriately enough, Nehemiah Grew made his name in plants and, in the interstices of a busy medical practice, anatomized the botanical kingdom; Regnier de Graaf busied himself with follicles and reproduction, mostly of rabbits, which labour reportedly so exhausted him that he died in his 32nd year; John Mayow did not live much longer, though he helped make the nitro-aerial spirit of his mentor Lower (see under 1691) more respectable among those inclined to philosophize.

1591 (4.0 centenary). Those who forgot to celebrate the birth of François Viète in 1590, when it enjoyed a 4.5¢ observance, can repair the omission this year, which marks the 400th anniversary of the publication of his *In artem analyticam isogage*. The book, which introduced a quasi-algebraic notation, including a short-hand designation of powers, marked an epoch in mathematics.

1551 (4.0 palindrome). This year appeared the *Prutenicae tabulae coelestium motuum*, calculated by Erasmus Reinhold (born 1511, 4.8¢) on the theories of Copernicus, in which he did not believe. The 'Prutenic tables' did much to ease the exploitation of Copernicus' system by mathematical astronomers. The first widely used tables based on new principles since the tables of King Alfonso (see 1221) and their derivatives, they served until replaced Kepler's (see 1641).

1541 (4.5 centenary). Reinhold received his first thorough account of Copernicus' theories in 1541, from his colleague Rheticus, then just back at Wittenberg, where he and Reinhold taught mathematics and astronomy, from a lengthy stay with Copernicus. A still more adventurous European, Hernando de Soto, penetrated to the Mississippi River, and another, Francisco de Orellane, discovered the Amazon.

Some lamented the death of the noisy doctor and hermetic chemist, Theophrastus Bombastus von Hohenheim, better known as Paracelsus.

1491 (5.0 centenary). A year cursed by five centuries of school children. It saw the publication of Filippo Calandri's *Aritmetica*, which presents long division in much the way that all of us not brought up with calculators learned it.

1441 (5.5 centenary, 5.0 palindrome). We expect something worthy of observance in this conjunction of half centuries and palindromes. Our best candidate, the *De docta ignorantia* of Nicholas of Cusa (born about 1401, 5.9¢), has against it only the date, which is usually given as 1440. But as Nicholas himself might have argued, from the principle of *coincidentia oppositorum* that underlies learned ignorance, he could be said to have completed his manuscript in the year after he finished

The guillotine (*Encyclopaedia Britannica*, 1800). showed that blood is made red by the 'nitrous spirit of air'.

1661 (3.0 palindrome). Boyle's *Sceptical Chymist*, published in 1661, consists of lengthy dialogues on the elementary principles of matter. The discussion arrives at the definition of a chemical element as a real substance not yet decomposed. In a companion piece of the same year, *Certain physiological essays*, Boyle discussed chemical combination in terms of the corpuscles of which he supposed all bodies to be compounded. Some see in these writings the beginnings of modern chemistry.

1641 (3.5 centenary). In an attempt to adapt the very fine to measurement of the very gross, William Gascoigne placed crosshairs movable by a micrometer screw in the focal plain of a keplerian telescope. The device was reinvented a little later by the French astronomers Auzout (see 1691) and Jean Picard, who incorporated it into useful instruments for surveying as well as for astronomy. Gascoigne did not long survive his brain child. He died in his early thirties, at the battle of Marston Moor. He thus joined, we suppose in the Heavens they both so assiduously studied, his younger contemporary Jeremiah Horrocks, who died in the sesquicentennial year 1641. Two years earlier, Horrocks had observed a transit of Venus that he had predicted from Kepler's 'Rudolphine

it. Learned ignorance teaches that the straight and the curved, the polygon and the circle, the centre and the circumference, the one and the many, coincide at infinity. The created world itself is relatively infinite; hence, strictly speaking, it cannot be said to have a centre; which, therefore, the Earth cannot occupy. Nor can it enjoy absolute rest, because

The mass production of bombs depended on the availability of plutonium, discovered by Edwin McMillan and Glenn Seaborg 50 years ago.

1941 was a good year for moulds. In Oxford, Howard Florey and his group were accelerating their work on what had just been first dubbed 'antibiotics'. At Stanford, George Beadle and Edward Tatum began irradiating the bread mould *Neurospora* and by carefully noting its altered biochemical capabilities, so initiated the train of research leading to their doctrine of 'one gene, one enzyme' and to a share of a post-war Nobel prize. Still in the United States, the German-born biochemist Fritz Albert Lipmann distinguished between high and low-energy phosphate bonds, of which ATP is metabolically the most important possessor of the former; and Charles Huggins (born 1901, 0.9¢) began treating prostatic cancers with hormones.

More scientists were alive in 1941 than had lived in all

The far side of the moon (NASA/Science Photo).

absolutes do not belong to creatures. Nicholas' universe was not Aristotle's.

1391 (6.0 centenary). "Litell Lowys my sone, I have perceived well by certeyne evidences thine abilite to lerne sciencez touchinge nombres & proporciouns; as well considere I thy bisy preyere in special to lerne the tretis of the astrelabie." Thus opens the little treatise on the astrolabe, subtitled "Bred and mylk for childeren", that Geoffrey Chaucer composed for his young son Lewis. The booklet describes the operation, but does not explain the construction, of the great medieval analogue computer. Although user-friendly, it is well beyond the capacity of most children in our decadent age.

1221 (7.0 palindrome). Grown-up accounts of astrolabes, together with translations of other Arabic works on astronomy, fill the *Libros del saber de astronomia*, commissioned by Alfonso X 'El sabio' of Castille, whose birth fell in the palindromic year 1221. Alfonso also sponsored an updated version of the Toledan tables for calculating planetary positions on Ptolemaic principles. The Alfonsine tables, in various forms, were standard among Western astronomers until the work of Erasmus Reinhold in the palindromic year 1551.

1941 (0.5 centenary). War dominated invention. The turbojet, patented by Frank Whipple in 1930, took wing. On the day before the attack on Pearl Harbor, President Roosevelt authorized the project to make an atomic bomb, which, with the jet plane, was to constitute the first 'weapons system' for mass annihilation.

past time. A proportionate number died. We can notice only a few: Henri Louis Bergson, the French philosopher; Annie Jump Cannon, the American astronomer; James George Fraser, the author of *The Golden Bough*; Arthur Erich Haas, the Austro-American physicist who took up his trade as a *pis aller* for a career as an historian of science; Sir Frederick Banting, of insulin fame; Tulio Levi-Civita, the Italian mathematician who exchanged insights into general relativity with Einstein; Dayton Clarence Miller, who laboured to subvert the special relativity; Hermann Walther Nernst, the German Nobel-prizewinning chemist who gave the world the Nernst lamp and the third law of thermodynamics; and French chemist Paul Sabatier, awarded a Nobel prize, for his way of hydrogenating organic compounds.

1941 ± 25 (0.25 and 0.75 centenary). The year 1966 belonged to the moon. The Soviet Union made the first soft landing there and also put the first lunar orbiter into its circle, accomplishments repeated by the Americans within a few months. Earthlings finally got a look at the far side of the moon.

Closer to home, Daniel Carleton Gajdusek succeeded in transferring kuru to chimpanzees, and so introduced the world to slow viruses. Another virus, that of German measles, received a body-blow, through the development, by Harry Meyer Jr, Paul Parman and J. C. Panos, of a live vaccine against it.

We record the deaths of Dutchman Luitzen Egbertus Jan Brouwer, the

founder of the intuitionist school of mathematics; his compatriot Peter Debye, who made important contributions to the old quantum theory of the atom; the cosmopolitan physical chemist Georg von Hevesy; the Belgian priest and cosmologist, Georges Lemaitre; and the Russian nuclear physicist Vladimir Iosifovich Veksler, the inventor, independently of E. W. McMillan, of the principle of phase stability in particle accelerators.

The Nobel prize for physics for 1966 went to Alfred Kastler for his ways of studying the interactions of radio waves with atoms. That for chemistry was awarded to Robert S. Mulliken for his work on chemical bonds and molecular orbitals. Charles Huggins shared his for physiology or medicine a quarter of a century (0.25¢) after his work was started.

Only literature merited a Nobel prize in 1916. Physics, chemistry and medicine were at work in the mechanized slaughter of the Great War. While a generation of young men were killing one another, the pacifist Albert Einstein published his complete and general theory of relativity. One of its consequences is that even light cannot escape from a star collapsed within a sufficiently small 'critical' compass. The connection between the war and the theory suggested by a black hole was embodied in Karl Schwarzschild, who worked out the critical size (the 'Schwarzschild radius'), and who perished from a disease contracted at the front. Another wartime casualty was Keith Lucas, a bright star of British physiology.

In 1916, Gilbert Newton Lewis, a still neutral American, proposed that the chemical bond consists of a pair of electrons shared by two atoms. It was the chemists' rather than the physicists' atom he described, but after the distractions of war his *Valence and the Structure of Atoms and Molecules* (1923) became a bible for a whole new generation of chemists.

Hiram Stevens Maxim, inventor of a popular machine gun, died quietly of old age. So did the historian-philosopher-physicist Pierre Duhem, who had spent the early war years demonstrating that the Germans had never done anything in science, and Ernst Mach, whose last reasoned contribution to science was a rejection of the theory of relativity; the American astronomer Percival Lowell, who interpreted the 'canals' of Mars as works of a civilization superior to our own; and the British chemist William Ramsay, discoverer of several inert gases, whose vicious verbal attacks on his former colleagues in Germany violated the spirit of his personal motto: 'Be kind'. □

J. L. Heilbron is at the Office for History of Science and Technology, University of California at Berkeley, 470 Stephens Hall, Berkeley, California 94720, USA; W. F. Bynum is at the Wellcome Institute for the History of Medicine, 183 Euston Road, London NW1 2BN, UK.

More on Silver Spring monkeys

SIR—G. Christopher Anderson's article, "Fight over Maryland monkeys" (*Nature* 343, 297; 1990) fails to provide appropriate information, context and background for the events and people involved.

Anderson describes the 17 monkeys in the "Silver Spring monkeys case" as having been "rescued" by police from inhumane conditions. This is colourful writing, but it is incorrect. Dr Taub was never convicted to keeping monkeys in inhumane conditions. He was explicitly acquitted of 118 of 119 original charges. The final charge, that of failing to provide necessary veterinary care to one (of 17) monkeys, was based on the contention that the animal had osteomyelitis that was the result of improper care. Galvin, the prosecuting attorney cited in Anderson's article, alluded to osteomyelitis several times in his closing remarks, despite his being aware that a pathology report clearly stated that the animal did not have osteomyelitis. After the appeal at which the one remaining charge against Taub was overturned, Galvin left the Maryland State's Attorney's office and became president of the Washington DC chapter of Lawyers for Animal Rights.

Thus, after five years in the courts, all the original 119 counts were dismissed. Animal rights groups invariably describe this as having been based on a technicality and Anderson presents this evaluation uncritically. The facts are otherwise. The decision of the court states that the Maryland Animal Cruelty Statute did not apply to federally funded laboratories because their practices (and specifically those of Taub) were already fully regulated and supervised; this included, but was not limited to, "periodic announced and unannounced inspections pursuant to the animal welfare act". Moreover, Taub has been further exonerated by four separate investigations.

What then happened in Taub's laboratory? Nine years ago, Alex Pacheco, an animal rights extremist, infiltrated Taub's laboratory as a "volunteer worker". During this time, Pacheco took photographs which several witnesses testified did not resemble conditions they had ever witnessed in Taub's laboratory. Pacheco's photograph of a monkey that he had helped to place in a primate-restraining chair has been used extensively for propaganda by People for the Ethical Treatment of Animals (PETA). PETA routinely acts as a publicist for the Animal Liberation Front, which is listed as a terrorist organization by the Federal Bureau of Investigation and Scotland Yard. PETA promptly receives copies of photographs, videotapes and documents stolen from laboratories during illegal break-ins by the ALF and other groups. Pacheco is co-

founder and director of PETA.

In the ensuing years, disputes over the custody and legal status of the monkeys have been exploited for publicity by PETA. The monkeys are at present housed at the Delta Regional Primate Center at Tulane University. When the monkeys were evaluated by panels of independent, distinguished veterinarians, killing was recommended for some of the monkeys whose health had deteriorated. PETA vigorously protested against this decision and secured a restraining order to prevent their being killed. The order was overturned in court and one of the monkeys was killed in January 1990. Three were killed in early July. Following a specially designed plan, investigators made electrophysiological measurements of neural activity in the brain while the monkeys were under anaesthesia just before being killed. Spinal cord and brain tissue were taken at autopsy. There is significant potential here for development of new treatments for victims of stroke and spinal cord injury. Nobel Laureate David Hubel, immediate past-president of the Society for Neuroscience, and many other distinguished neuroscientists have attested to the great value of the unique information gained from investigations of the brains of monkeys with limbs that had been deafferented for many years.

After the death of the first monkey, animal rights activists vowed to return to the courts to prevent the killing of any of the survivors, and the Physicians' Committee for Responsible Medicine (PCRM) filed a complaint with the Department of Health and Human Services. Dr Neal Barnard, quoted by Anderson, is a Washington, DC, psychiatrist who opposes the use of animals in research. He is a medical adviser to PETA and heads PCRM. At its June 1990 meeting, the American Medical Association publicly criticized PCRM for misrepresenting the critical role animals play in research. The AMA reaffirmed the overwhelming support of physicians for the need for continuation of research.

The activists' preference for prolonging the lives of animals in failing health will ensure live material for several more years of publicity but can be of little comfort to the monkeys. The "saga of the Silver Spring monkeys" has been widely used to attack all biomedical research using animals. PETA and PCRM are not unbiased sources of information. Your readers deserve a more critical review of the issue.

S. B. GALSWORTHY

Department of Microbiology & Immunology,
University of Western Ontario,
Health Sciences Centre,
London, Canada N6A 5C1

Europe's future

SIR—When European governments "conduct themselves in a seemly fashion" (*Nature* 348, 267; 1990), not only do they promote the ideals of Western civilization, but they also foster a much needed spirit of decency and hope in countries which, to this day, suffer the indignation and suppression of kaiserian or stalinistic acts.

The interdependence of European peoples expressing common ideals and actions will become a social phenomenon. If we fault their chief actors for lack of imagination in Paris, we must be inspired by the persistence, faith and joy of Prague and the peoples of Eastern Europe as a whole.

The three Baltic countries of Estonia, Latvia and Lithuania were more than "nominally independent states" a few years ago, they were full members of the League of Nations until the Second World War left them occupied, much as Kuwait is today. It is noteworthy that in Paris, both the Soviet parliament of Estonia and the National Congress of Estonia (elected by citizens of the Republic of Estonia) had sent delegations, but these were rejected by their seemingly "civilized" partners.

I, too, look forward to the convening of the real Congress of Paris.

ILLIMAR ALTOSAAR

Faculty of Medicine,
University of Ottawa,
40 Marie Curie Street,
Ottawa, Ontario, K1G 1X2, Canada

Hunger strike

SIR—In connection with the letter "Hunger strike" from V. Gurzadyan (*Nature* 347, 610; 1990) on the hunger protest of Professor Victor Ambartsumian against violations of the human rights of Armenian inhabitants in the mountainous Karabagh region of Azerbaijan, the Committee of Astronomy of the Polish Academy of Sciences, representing the astronomical community of Poland, declares its full solidarity with the oppressed Armenians and expresses its deepest concern about the health of Professor Ambartsumian, one of the most distinguished scientists of this century and former president of the International Astronomical Union and of the International Council of Scientific Unions.

The Committee of Astronomy appeals to the astronomical institutions of the Soviet Union and to all international scientific organizations to take all possible steps in support of the protest of Professor Ambartsumian.

ROBERT GLEBOCKI
(Chairman)

Committee of Astronomy,
Polish Academy of Sciences,
Warsaw, Poland

Cold dark matter makes an exit

David Lindley

WHEN, about two decades ago, cosmologists began to apply statistical analysis to the distribution of galaxies in the sky, one of the first results was a schism in the theoretical world. The article of faith on which this schism hinged was the existence or not of structures in the distribution of galaxies — superclusters, filaments, giant voids — on very large scales. Over the years, evidence has mounted that big structures really do exist, and the paper by Saunders *et al.* in this week's issue¹ seems finally to prove the point — and in so doing, disposes of the favoured 'cold dark matter' model of galaxy formation.

The statisticians discovered early on that the two-point correlation function, which measures the difference between the observed galaxy distribution and a strictly random one, diminishes smoothly for increasing galaxy separation, suggesting that whatever mechanism formed galaxies operated on small scales, but petered out on large scales, so that any large-scale structures were merely the predictable aggregates of the small-scale clustering. But several observers obstinately insisted that the filaments and voids they could see plainly in galaxy maps were more than the two-point correlation function could account for.

Theoretical preference

It was hard, however, for the disbelievers to come up with any unambiguous statistics with which to prove that *bona fide* large-scale structures existed. Theoretical preferences hinged on this observational debate. In 'bottom up' theories of galaxy formation, structures formed first on small scales and aggregated into larger units; in 'top down' theories, the large structures were primary, and fragmented into small pieces.

This debate now seems, in hindsight, rather pointless. As cosmologists attacked the galaxy distribution, it became apparent that the existence of a smooth correlation function on small scales does not preclude the existence of large-scale structures (this is only a fancy way of saying that an infinite number of statistical distributions can have the same mean and standard deviation but have very different higher-order moments).

Equally, evidence mounted that large structures really do exist. Recently we have had the 'Great Attractor'² a giant congregation of galaxies and clusters noticeable by the gravitational pull it exerts on its neighbouring galaxies, the 'Great Wall'³, a continuous sheet of galaxies stretching across the sky and, with less certainty, periodicity in the galaxy distribution on a 100-megaparsec scale⁴.

The problem is then to produce a model which can account for the small-scale structures, characterized as ever by the two-point correlation function, but also permits the possibility of large-scale structures. The disparate nature of the various large structures means that one should not hope to predict them in any specific way: it is sufficient to show that one's theory harbours some reasonable probability that such a structure might arise in the volume of the observable universe.

For a few years, the best attempt at such a comprehensive theory appeared to be the cold dark matter (CDM) model. Dark matter is the invisible stuff that dynamical studies of galaxies and clusters of galaxies indicate must be there, but which can't be seen; the fact that it is cold means that the particles which compose it (entirely hypothetical particles, it must be said) are slowly moving, a property that enables them to clump together into galaxy-sized lumps.

But now Saunders *et al.*¹ say that the CDM theory, like many of its predecessors, must be discarded. They argue that the picture of the galaxy distribution given by the Infrared Astronomy Satellite (IRAS) is, on the largest scales, in clear contradiction with what CDM would have. This disavowal of CDM is all the more remarkable for coming from a group of authors that includes some of the theory's long-time supporters.

Saunders *et al.* can make this definitive statement because they avoid arguments about whether CDM or any other theory can explain such particular observational features as the Great Attractor or the Great Wall. Both the Greats are clearly unique aggregations, and although their existence in CDM or any other theory may be a little unlikely, cosmologists could carry on believing their favourite theory provided they were willing to accept a couple of low-probability items in their version of the heavens. But the IRAS survey allows the well-studied correlation-function technique to be applied on the scale of the whole sky. Infrared observations avoid most of the extinction problems which limit the sky-coverage and uniformity of a comparable optical survey, and by producing a survey that covers almost the whole sky to uniform depth, IRAS generates a catalogue of galaxies from which one can obtain, for the first time, a believable correlation function over a very large range of scales.

The CDM model slips up because it does not get the galaxy distribution right on scales of about 20 megaparsecs and above: the real sky has more non-randomness than CDM can provide. The reason CDM fails is essentially that it was

designed to explain structure at the level of galaxies and small clusters of galaxies. In the early days of dark-matter cosmologies, the idea that neutrinos might have a cosmologically interesting mass was in vogue, and although neutrinos may indeed have a mass that will interest cosmologists^{5,6}, the fact that they must be moving relativistically makes it almost impossible for them to aggregate in gravitationally bound associations on the scale of individual galaxies. CDM, by contrast, works on the galactic scale. Large scales were never imagined to be its forte, but only with the work of Saunders *et al.* has quantitative evidence of this inadequacy come to light.

Cosmological constant

Where does this leave galaxy-formation theory, and cosmology in general? Some of Saunders's colleagues have suggested that CDM can be saved, at least in modified form, if a non-zero cosmological constant is resurrected^{7,8}; a cosmological constant, amounting to a vacuum energy density, modifies cosmic dynamics on the large scale, but leaves the small scale alone, compensating for CDM's weakness. Alternatively, perhaps neutrinos really do have a cosmologically significant mass^{5,6}, so that one can imagine the dark matter being mostly CDM for individual galaxies and mostly neutrinos for the giant structures.

If cosmologists are forced to start throwing around several kinds of dark matter, along with a cosmological constant, scepticism is bound to arise: a sufficiently complicated model can always replicate, in the manner of Ptolemy's epicycles, a limited set of observations.

Ptolemy's Solar System was satisfactorily replaced by one dictated by a single underlying principle, the inverse-square law of gravitation. But do we have any right to expect so messy a subject as galaxy formation to be likewise guided by a single idea? If particle physicists prove tomorrow that the neutrino has a mass of 4.73 electronvolts, and that a species of previously unknown particle provides 0.62 of a critical cosmic density in the form of cold dark matter, cosmologists could confidently set themselves to constructing messy theories with these ingredients. But if cosmologists declared, in the absence of independent evidence, that only with these same ingredients could they explain the distribution of galaxies, would anyone believe them? □

David Lindley is Associate Editor of *Nature*.

1. Saunders, W. *et al.* *Nature* **349**, 32–38 (1991).
2. Lynden-Bell, D. *et al.* *Astrophys. J.* **326**, 19–49 (1988).
3. Geller, M. & Huchra, J. *Science* **246**, 897–903 (1990).
4. Broadhurst, T.J. *et al.* *Nature* **343**, 726–728 (1990).
5. Sciama, D.W. *Nature* **348**, 617–618 (1990).
6. Maddox, J. *Nature* **348**, 579 (1990).
7. Efstathiou, G., Sutherland, W.J. & Maddox, S.J. *Nature* **348**, 705–707 (1990).
8. Bertschinger, E. *Nature* **348**, 675–676 (1990).

The essence of inactivity

Kevin Davies

ALTHOUGH 30 years have passed since Mary Lyon proposed the theory of mammalian X-chromosome inactivation¹, rather less is known about the rule that leads to dosage compensation than about the handful of genes that are exceptions to it. Three papers²⁻⁴, two on pages 38 and 82 of this issue^{2,3}, describe the characterization of two genes that will considerably advance our understanding of X inactivation. Both genes map to the long arm of the human X chromosome (Xq13). The first exhibits a unique pattern of expression and is close to, and perhaps the same as, the X-inactivation centre (*XIC*)^{2,3}; the other is the first gene from the long arm found to 'escape' X inactivation, and halved dosage (haploinsufficiency) of this gene may give rise to some or all of the features of Turner's syndrome⁴.

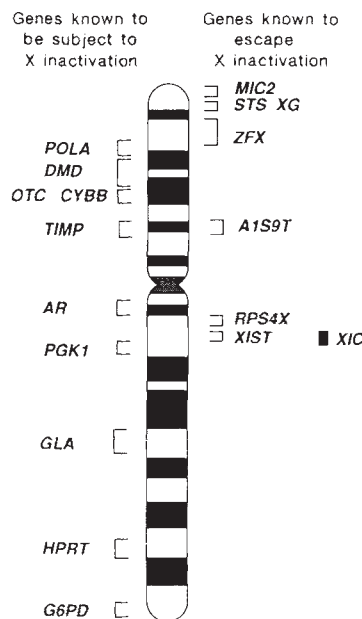
Random inactivation of one of the two X chromosomes in cells of placental mammalian (XX) females during early development ensures that they have the same dose of X chromosome genes as males possessing only one X chromosome. Studies of hybrid cell lines have loosely mapped the *XIC* to the proximal region of Xq, from where an inactivating 'signal' is conventionally thought to radiate. This signal was formerly envisaged to shut off transcription of all genes except those at the distal end of the short arm (Xp22.3) — *MIC2* and *XG*, both encoding cell-surface proteins, and the steroid sulphatase gene (*STS*).

Over the past year or so, however, the list of genes that escape inactivation has grown to include *ZFX* (ref. 5), which maps to Xp21.3-22.1 (ref. 6), *AIS9T* (which complements a murine DNA synthesis mutation) at Xp11 (ref. 7) and most recently *RPS4X* on the long arm (see figure)⁴. So inactivated genes are interspersed with genes that escape inactivation, implying that the former are shut down not by virtue of their regional localization, but on a locus-by-locus basis. But the mechanism by which this may be achieved has remained a mystery.

The latest work from Hunt Willard's group^{2,3} precisely localizes *XIC* to Xq13 (ref. 3), and identifies a gene within the same interval — *XIST* (X inactive-specific transcripts) — whose unique expression from the inactive, but not the active, X chromosome suggests that it has a role in the induction of X inactivation². *XIST* was isolated, fortuitously it seems, during efforts to clone homologues of *STS* (ref. 8) by antibody screening, but DNA sequencing has yet to reveal homology to any known gene — including *STS*. Although the *XIST* RNA is polyadenylated, the presence of many stop codons in the exons

sequenced so far suggests that *XIST* may encode not a protein but a structural RNA, which Brown *et al.* contend is easier to formulate into models of X inactivation. But how the one active X chromosome is distinguished from the inactive copy is unclear.

Knowledge of the complete gene structure of *XIST* will probably be essential to establish its precise role in X inactivation,



Map of the human X chromosome depicting genes and their inactivation state (adapted from ref. 7).

which Brown *et al.* acknowledge could be a secondary one. Attention will turn towards the murine homologue to *XIST* (*Xist*), and in particular to whether or not it maps to the conserved region in mice, *Xce* (X chromosome-controlling element)⁹. Another priority will be to examine the early expression of *Xist*, because X chromosomes are inactivated within 3-6 days of murine development. But human geneticists should not lose heart. Rare families manifesting female-to-female transmission of X-linked traits, such as haemophilia B (ref. 10), may be indicative of a co-inherited defect in the *XIC*, resulting in the exclusive inactivation of the normal chromosome. This may in turn signify an abnormality in *XIST*.

As *XIST* is only expressed in cell lines containing one or more inactive X chromosomes, and is therefore female-specific in karyotypically normal individuals, *XIST* may be involved in the phenotype of X-chromosome disorders such as Klinefelter's and Turner's syndromes². This might also be true for other genes that escape X inactivation³, and an intriguing candidate for Turner's syn-

drome is a gene recently cloned in David Page's laboratory by Fisher *et al.*⁴.

Turner's syndrome, which results from a common sex chromosome disorder, is characterized by short stature, somatic abnormalities such as webbed neck, ovarian degeneration and a 99 per cent death rate *in utero*. It is usually associated with the loss of a sex chromosome (45,X genotype) and is thought to be a consequence of single dosage (monosomy) of a gene or genes present on both X and Y chromosomes¹¹. From studies of X-chromosome rearrangements it seems that more than one gene on the X chromosome may contribute to the various features of Turner's syndrome¹¹⁻¹³, although their locations are poorly defined. But deletion mapping in 46,XY Turner's females has refined one probable Y chromosome localization to a 90-kilobase stretch⁴, between the candidate sex-determining gene, *SRY* (ref. 14), and more proximally, *ZFY* (ref. 15).

Fisher *et al.* identified the candidate gene by searching for transcribed sequences on the Y chromosome within the 90-kilobase interval, and found just one — *RPS4Y*. A related complementary DNA turned out to encode *RPS4X*, which is just proximal to *XIST* in Xq13. *RPS4X* is the only gene from the long arm known to be expressed from both active and inactive X chromosomes⁴ and has a homologue on the Y chromosome. Both are criteria that must be satisfied by candidate genes for Turner's syndrome.

The X- and Y-encoded *RPS4* genes, which are 93 per cent identical to each other, reveal maximal homology to the gene encoding the rat ribosomal S4 protein¹⁶, a constituent of the small ribosomal subunit. (A gene probably identical to *RPS4X*, called *SCAR*, was cloned previously¹⁷ but no sequence similarity was appreciated at the time.) Although by no means proof of the role of *RPS4X* and *RPS4Y* in Turner's syndrome, haploinsufficiency of a ribosomal protein might be expected to have profound effects

1. Lyon, M. F. *Nature* **190**, 372-373 (1961).
2. Brown, C. J. *et al.* *Nature* **349**, 38-43 (1991).
3. Brown, C. J. *et al.* *Nature* **349**, 82-84 (1991).
4. Fisher, E. M. C. *et al.* *Cell* **61**, 1205-1218 (1990).
5. Schneider-Gadicke, A., Beer-Romero, P., Brown, L. G., Nussbaum, R. & Page, D. C. *Cell* **57**, 1247-1258 (1989).
6. Page, D. C. *et al.* *Genomics* **7**, 37-46 (1990).
7. Brown, C. J. & Willard, H. F. *Am. J. hum. Genet.* **46**, 273-279 (1990).
8. Ballabio, A. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **84**, 4519-4523 (1987).
9. Keer, J. T. *et al.* *Genomics* **7**, 566-572 (1990).
10. Taylor, S. A. M., Deugau, K. V. & Lillicrap, D. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
11. Epstein, C. J. in *Turner Syndrome* (eds Rosenfeld, R. G. & Grumbach, M. M.) 13-28 (Dekker New York, 1988).
12. Therman, E., Laxova, R. & Susman, B. *Hum. Genet.* **85**, 455-461 (1990).
13. Therman, E. & Susman, B. *Hum. Genet.* **85**, 175-183 (1990).
14. Sinclair, A. H. *et al.* *Nature* **346**, 240-244 (1990).
15. Page, D. C. *et al.* *Cell* **51**, 1091-1104 (1987).
16. Devi, K. R. G., Chan, Y.-L. & Wool, I. G. *Biochim. biophys. Acta* **1008**, 258-262 (1989).
17. Wiles, M. V., Alexander, C. M. & Goodfellow, P. N. *Somat. cell. mol. Genet.* **14**, 31-39 (1988).
18. Kongsuwan, K. *et al.* *Nature* **317**, 555-558 (1985).

on cell growth. Indeed, some *minute* mutants of *Drosophila*, which are characterized by anatomical abnormalities with variable lethality and fertility, represent defects of ribosomal protein genes¹⁸.

The weak link in contemplating a role for the *RPS4* loci in Turner's syndrome is the position of *RPS4X* on Xq. Most evidence supports a localization on Xp, as short-arm deletions generally produce the most severe combination of Turner's

phenotypes¹¹. One possible explanation to resolve this problem is that the Turner's gene(s) may not be deleted but, in some cases, simply inactivated by position effects on structurally abnormal X chromosomes¹³. Which brings us back, perhaps, to *XIST*. Now that its existence is known, future investigations into this gene promise to be anything but inactive. □

Kevin Davies is an assistant editor of *Nature*.

PALAEOMAGNETISM

Plate tectonic speed limits

Richard G. Gordon

WHEN W. Alvarez and colleagues¹ presented results from a palaeomagnetic study of limestone blocks from the Franciscan complex near Laytonville, California, the results were greeted with scepticism by many geologists. After all, the palaeomagnetic samples come from two blocks of deep-ocean limestone lying within a tectonically dismembered *mélange* — rocks so stirred up during the process of transport from the deep sea to accretion in the California coast ranges that little, if any, of the original positional and attitudinal relationships between rock layers is preserved. However, what really prompted the scepticism was not these weaknesses, but the results — inferred plate speeds far exceeding the speed at which any plate has moved in the geologically recent past. Now, from a greatly expanded and detailed palaeomagnetic study of limestone blocks from the same formation, Tarduno *et al.*² have laid to rest most, if not all, of the original objections to the interpretation of very rapid plate motion.

The key data in these analyses are palaeolatitudes, which are inferred by simple calculation from the inclination, relative to the palaeohorizontal, of the magnetization vector preserved in rocks (Fig. 1). To know whether the inclination was originally positive (downward) or negative (upward), one must know which way is up in the rocks. This ambiguity was resolved by palaeontology combined with the original sedimentary layering preserved in the blocks. From changes in the character of the foraminifera, I. Premoli-Silva, a collaborator in the original study¹ was able to determine the direction in which each block became younger, which is of course the ancient upward direction.

For most palaeomagnetic results the hemisphere of origin would still be ambiguous, however, because of geomagnetic reversals. For example, rocks magnetized in the northern hemisphere during the past 700,000 years would have a positive inclination, whereas rocks magnetized at an equal southern latitude would have an equal negative inclination.

During times of reversed geomagnetic polarity, however, the signs of these inclinations would be reversed. Fortunately all the rocks of the Laytonville Limestone were deposited (and presumably magnetized) during the longest known interval of uninterrupted normal magnetic polarity, which persisted without a geomagnetic reversal from about 85–116 million years ago. Thus the magnetic polarity of the rocks is known and their mainly negative inclinations show that the rocks originated in the southern hemisphere, far away from their present location.

From the large implied distance of travel, the age of the rocks and the inferred age of their accretion onto California, the northwards component of plate velocity can be estimated. The northwards component so estimated by Alvarez *et al.*¹ was astonishingly fast — 25 cm yr⁻¹. Of course, it is unlikely that the plate travelled precisely due north. From reasonable estimates of the direction of convergence between the oceanic plate carrying the limestone and the North

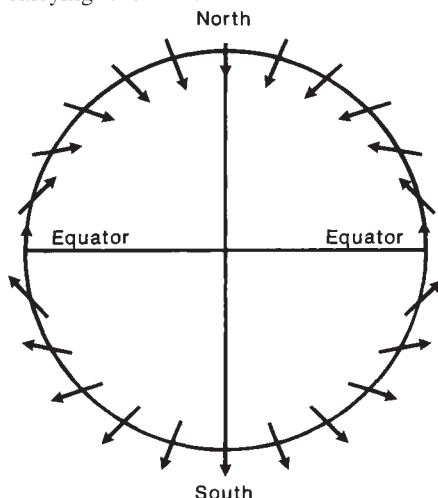


FIG. 1 Magnetic inclination and latitude. The sketch shows the inclination of the time-averaged geomagnetic field during an interval of normal magnetic polarity as seen by a distant observer lying in Earth's equatorial plane. During intervals when the geomagnetic field was reversed, the time-averaged field would be opposite to the directions shown here.

American plate, Alvarez *et al.* estimated a convergence rate of 38 cm yr⁻¹, 3 to 4 times faster than the fastest present convergence rate³ of about 11 cm yr⁻¹.

Tarduno *et al.*² confirm the southern-hemispheric origin of the Laytonville Limestone while extending the study from two to nine blocks. The palaeontology indicates that five of these blocks are overturned relative to their original orientation whereas the remaining four blocks are right side up. When the magnetization vectors are corrected to their original orientations, the results from different blocks come into much better agreement than when examined in their present orientations. The results strongly support the contention that the blocks travelled rapidly between 90 and 50 million years ago on the way to their present locations. The many samples span enough geological time for Tarduno *et al.* to estimate the northwards component of plate velocity not only after, but during, deposition of the limestone. They find the northwards velocity to be 28 cm yr⁻¹ with all data included and a slower but nevertheless very fast northwards component of 15 cm yr⁻¹ when outlying data are omitted.

As I expect is true for many other geologists, my first reaction to the new results was to wonder what could be wrong with the data. With enough searching, weaknesses could be found, as they could with any data set. One possible weakness is in the large uncertainty in converting palaeontologically determined ages to ages in millions of years. Another is the hard-to-explain scatter of a few of the data.

But why is it so hard to accept very rapid ancient plate velocities? The answer is that plates do not move so rapidly today, and one of the multiple meanings of the principle of uniformity tells us that rates of geological processes must be the same in the past as they are today. This interpretation of uniformitarianism is demonstrably wrong⁴. We should instead apply an alternative meaning — that the laws of physics and chemistry apply as well to the past as to the present.

Consider the case for a speed limit to plate velocities. In an important paper, Forsyth and Uyeda⁵ proposed that all fast-moving plates moved at a nearly uniform speed reflecting the balance between two dominant forces, that due to the pull of the negative buoyancy of subducting slabs attached to plates, and that due to the viscous resistance to the penetration of the slabs into the mantle, leading to the concept of a terminal velocity of subduction (Fig. 2). In their model they assumed that the negative buoyancy of subducting slabs did not differ along a trench or between trenches and that other characteristics of the surface portion of the plate, including its area, are unimportant to the balance of forces. If this model is accepted without modification, then plates could

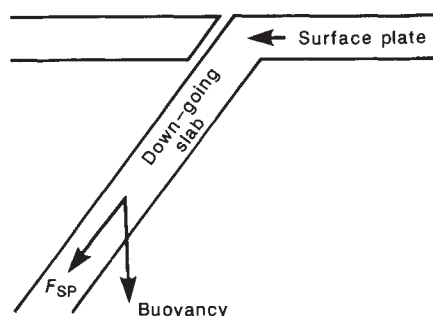


FIG. 2 Forces in plate tectonics: a simplified cross-section of a cold, dense, subducting lithospheric slab attached to a surface lithospheric plate. The lithosphere acts as a stress guide, transmitting the slab pull (F_{sp}) — the down-dip component of the 'negative buoyancy' of the down-going slab — to the surface plate to drive its motion.

not have moved any faster in the past than they do today.

But this model can be challenged for several reasons. The negative buoyancy of subducting slabs should increase with the age of the seafloor delivered to trenches, owing to the cooling, subsidence and thickening of the lithosphere⁶. Similarly, a driving force that is due to the subsidence and thickening of the lithosphere with age and is distributed across the entire surface plate (although sometimes referred to as ridge push), should also increase with the age of the seafloor delivered to trenches⁷. In models with an age dependence of these forces, other forces including the drag at the base of plates (proportional to plate area) are found to be important⁸. Moreover, estimates of the viscosity of the subasthenospheric mantle have decreased since Forsyth and Uyeda's paper was written: slab resistance is now thought to be negligible.

In this light, what might be the cause of fast plate motion in the past and why does it not operate now? There are now only about 12 major plates, so few that it is doubtful that all possible plate geometries are represented⁹. One configuration missing from the catalogue of present plates is a plate with small area but delivering old, cold and dense lithosphere to a trench¹⁰. Tarduno *et al.*² argue that the Laytonville Limestone was deposited on such a plate, now entirely subducted beneath North America. Other similarly configured plates may have existed in the past, but because they are eventually subducted entirely, there is little record preserved.

Alternative interpretations of the new data are possible. For example, Hagstrum¹¹ suggests the rocks were remagnetized late in their history as they started to subduct along with the underlying plate. This alternative may be correct, but if there was fast northwards motion during deposition of the Laytonville Limestone (as indicated by the new data), in agreement with the sense and approximate rate of motion during later transport, then outright rejection of the hypothesis sug-

gested by Alvarez *et al.* becomes difficult.

Tarduno *et al.*² have squeezed out all of the palaeomagnetic information available from the Laytonville Limestone, but I think the problem they tackle has yet to reach the end of the research line. Complementary analyses of Pacific-plate palaeomagnetism and Pacific-basin plate motion are needed for further tests and for rigorous estimates on the speed limits of oceanic plates. Plate motions over the past 70 million years are reasonably well constrained. Available reconstructions for earlier times are useful working models, but can and should be improved through new data and improved methods of analysis.

Once again, palaeomagnetism has yielded surprising results that are probably of central geodynamic importance. If these results and interpretations withstand the test of time, they will provide strong constraints on plate dynamics and

EMBRYOLOGY

Molecule of the moment

Jonathan Slack

ACTIVIN is the molecule of the moment for embryologists. Fifty one years ago came the first report of a mesoderm-inducing factor that causes ectoderm in early amphibian embryos to differentiate into mesoderm tissues¹, and in the past few months mesoderm-inducing factors from a variety of sources have all been identified as activin A (refs 2–4). Even a shadowy substance, long known as the 'vegetalizing factor', has finally been purified and although no sequence is yet available it too looks like activin⁵. Unlike the previously identified mesoderm-inducing factors belonging to the fibroblast growth factor group, activins induce axial mesoderm (notochord and segmented myotomes) at high concentration as well as ventral-type mesoderm at low concentration.

This work all shows that activin can induce the axis, but does it actually do so in the early embryo? Thomsen *et al.*⁴ looked at the expression patterns of activin A and B, which are closely related gene products, in the early *Xenopus* embryo. They suggest that activin B is more likely to be a natural mesoderm-inducing factor because its expression commences in the late blastula, whereas activin A is not significantly expressed until the early tadpole stage. But even the late blastula is probably too late a stage for mesoderm-inducing factor to start operating, in that studies of vegetal cells suggest that it is emitted as early as the 64-cell stage⁶. This is two to three hours before the zygotic genome begins to be transcribed, so it is generally thought that at least the messenger RNA for the inducer,

on the evolution of the Pacific basin and its margin. If they do not, it will be because they have stimulated research in palaeomagnetism and Pacific basin plate motions along worthwhile directions. □

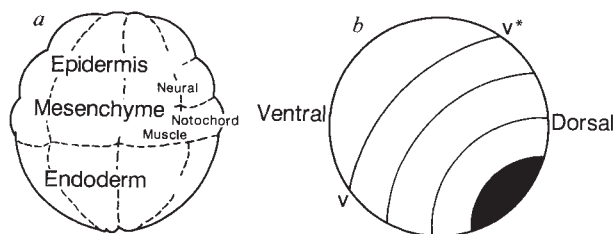
Richard G. Gordon is in the Department of Geological Sciences, Northwestern University, Evanston, Illinois 60208-2150, USA.

1. Alvarez, W., Kent, D.V., Premoli-Silva, I., Schweickert, R.A. & Larson, R.A. *Geol. Soc. Am. Bull.* **91**, 476–484 (1980).
2. Tarduno, J.A., McWilliams, M.O. & Sleep, N. *J. geophys. Res.* **95**, 15503–15527 (1990).
3. DeMets, C., Gordon, R.G., Argus, D.F. & Stein, S. *Geophys. J. Int.* **101**, 425–478 (1990).
4. Gould, S.J. *J. geol. Edu.* **15**, 149–150 (1967).
5. Forsyth, D.W. & Uyeda, S. *Geophys. J. R. astr. Soc.* **43**, 163–200 (1975).
6. Harper, J.F. *Geophys. J. R. astr. Soc.* **55**, 87–110 (1978).
7. Lister, C.R.B. *Nature* **257**, 663–665 (1975).
8. Hager, B.H. & O'Connell, R.J. *J. geophys. Res.* **86**, 4843–4867 (1981).
9. Solomon, S.C., Sleep, N.H. & Jurdy, D.M. *J. geophys. Res.* **82**, 203–212 (1977).
10. Schult, F.R. & Gordon, R.G. *J. geophys. Res.* **89**, 1789–1800 (1984).
11. Hagstrum, J.T. *Tectonics* **9**, 1221–1233 (1990).

and perhaps the protein as well, is maternal in origin and already present in the oocyte.

How can we prove that a substance really is an inducing factor *in vivo*? Obviously, it must be active and it must be present in the signalling region, in sufficient quantity, at the appropriate embryonic stage. It also has to be shown that the factor is actually secreted by the signalling cells and is transmitted to the responding cells. Activin A and B have passed few of these tests. They are both active, but are probably not expressed early enough in development, and their regional distribution in the embryo is as yet unknown. There is so far no evidence that the factor (or factors) secreted by the vegetal cells is really activin. In August, at a meeting at Les Diablerets, Switzerland, Thomsen, Melton and colleagues reported that the *Xenopus* oocyte contains several messenger RNAs similar to those of the transforming growth factor type β family, and some of these will probably show mesoderm-inducing activity. So if the activins do not make it to the finishing line there are other potential runners.

It is in fact difficult to understand how one substance alone can be responsible for mesoderm induction. If a substance spreads out by any diffusion process from a centre in the dorsovegetal quadrant then contours of equal concentration would run right around the egg (see figure). Whatever the state of specification of tissue on the ventral side (v) would be the state on the dorsal side, but further towards the animal pole (v*). So if we assayed the specification of small tissue



a, Specification map of the early *Xenopus* embryo, showing the tissue types formed by explants from different regions cultured in a neutral medium. b, What a specification map would look like if it were controlled by a single substance diffusing from the dorsovegetal region.

regions by culture in isolation we would see ventral-type differentiation on the dorsal side. This is not evident even in the most careful and discriminating studies^{7,8}, and the prediction represents the *reductio ad absurdum* of the one-signal model.

In the past few weeks the activin race-track has expanded to embrace the early chick embryo. Mitrani and coworkers^{9,10} have demonstrated that an isolated epiblast (the upper layer of the two-layered early embryo) will form axial mesodermal structures when incubated in activin, and they also claim to show the presence of activin B mRNA in the hypoblast (the lower layer)¹⁰. The data on which the second claim is based are rather weak, but whether it is correct or not it is important to realize that the biology of the early chick is much less well understood than that of *Xenopus*.

The conventional textbook story is that in the chick the primitive streak, which later forms the axial mesoderm, is induced from the epiblast by the underlying hypoblast. This view is based on microsurgical experiments¹¹ and may indeed be correct, but the critical experiments did not use cell labels so it cannot be guaranteed that the results are not due to cell movement rather than induction. It does seem, from cell-labelling work¹², that all the important events happen in a very small region at the posterior edge of the embryo (the posterior marginal zone) and that if induction is occurring it must take place at a rather earlier stage than previously thought. The isolated chick epiblast is also not as good a test tissue as the ectoderm of an amphibian embryo. In the absence of treatment, the *Xenopus* animal cap forms a solid ball of epidermis, but the chick epiblast forms various tissues including mesenchyme and

blood cells, and sometimes muscle as well, all of which are indisputably mesodermal in character. This means that the positive outcome following treatment is the formation of axial mesodermal structures rather than mesoderm as such; and because it is hard to achieve a good degree of terminal differentiation in chick epiblast

cultures, there is inevitably a higher degree of subjectivity in scoring the results.

SOLAR SYSTEM

Small grains of truth

Joe Nuth

THE origin of the Solar System from a rotating disk of gas and dust is becoming well understood — or so it is generally believed. But how that solar nebula formed from the pre-existing clouds of interstellar material is less clear. Isotopic data from the oldest Solar System materials presented by Zinner *et al.* on page 51 of this issue¹ are vital clues to answering this question.

Reports of isotopic anomalies in primitive Solar System materials have become almost commonplace recently thanks to the remarkable advances in analytical instrumentation and chemical separation procedures which have occurred over the past several years. Anomalies in noble gases served as tracers of a variety of grains that originated from other stars as they were isolated from chondritic meteorites. Zinner *et al.* now report studies of the trace-element compositions of circumstellar graphite and silicon carbide which reveal extinct radioactive ²⁶Al (in the form of its stable daughter ²⁶Mg) at concentrations (compared to ²⁷Al) more than 4,000 times higher in some grains than previously observed in Solar System materials.

The authors also demonstrate that a good correlation exists between the concentrations of nitrogen and aluminium in these grains. Such a correlation may indicate that the aluminium condensed as AlN in the outflow from carbon-rich stars and that AlN may have served as a condensation site for the formation of graphite and SiC grains. This latest study¹ also supports recent work by Lewis, Amari and Anders² suggesting that the SiC grains separated from primitive meteorites had to have come from several different stellar sources. In the case of the SiC grains, the most probable source³ is condensation in the outflows from stars on the asymptotic giant branch (AGB) of the main sequence of stellar evolution. Although AGB stars may have produced circumstellar graphite

All of us who work on the early development of *Xenopus* would like the chick system to be similar because it would increase the significance of our own work. But we really require some better embryology to assess the role of activin — in particular we need solid evidence that the streak is really induced by the hypoblast, and we need a better test tissue than the isolated epiblast □

Jonathan Slack is in the Developmental Biology Unit, ICRF Laboratories, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.

grains as well, it is more likely that the graphite grains found in primitive meteorites were produced in novae¹.

The high ²⁶Al/²⁷Al ratios observed in known circumstellar condensates might provide support for one model for the initiation of the collapse of the solar nebula proposed by A.G.W. Cameron (Harvard University) at a recent workshop*. In this model, the wind from an AGB star triggers the collapse of nearby regions of a giant molecular cloud and mixes live radioactive nuclei into the protostar. As these protostellar nebulae evolve, bipolar outflows from the protostars trigger the collapse of other portions of the cloud which, in turn, evolve and trigger new waves of star formation. During each wave of star formation the original ²⁶Al from the AGB progenitor is diluted by pre-existing interstellar material and lost through radioactive decay. The highest ²⁶Al/²⁷Al ratio observed in material known to have been processed in the primitive solar nebula is about 5×10^{-5} whereas the ratio expected in the first generation of stars triggered by an AGB wind is approximately 10^{-3} . On these grounds, Cameron concludes that our Sun formed as a second- or third-generation star in this hypothetical chain of events.

If this scheme is correct, it is possible that the SiC and graphite grains containing high ²⁶Al/²⁷Al may have passed through one or more protostellar systems before incorporation into our own. Sequential passage through several protostellar systems could provide a natural explanation to some of the anomalously low pre-solar cosmic-ray exposure ages found for SiC grains by Anders and coworkers³. In this model the population of chemically resistant, young, gas-poor grains inferred by their work could have been freshly synthesized in the AGB outflow while older, gas-rich grains could have been

1. Chuang, H. H. *Wilhelm Roux Arch. EntwMech. Org.* **139**, 556–638 (1939).
2. Smith, J. C., Price, B.M.J., Van Nimmen, K. & Huylebroeck, D. *Nature* **345**, 729–731 (1990).
3. Albano, R. M. *et al. Development* **110**, 435–443 (1990).
4. Thomsen, G. *et al. Cell* **63**, 485–493 (1990).
5. Plessow, S. *et al. Cell Diffn Dev.* **32**, 27–38 (1990).
6. Jones, E. A. & Woodland, H. R. *Development* **101**, 557–563 (1987).
7. Holtfreter-Ban, H. thesis (Univ. Rochester, 1965).
8. Kanéda, T. & Hama, T. *Wilhelm Roux Arch. dev. Biol.* **187**, 25–34 (1979).
9. Mitrani, E. & Shimoni, Y. *Science* **247**, 1092–1094 (1990).
10. Mitrani, E. *et al. Cell* **63**, 495–501 (1990).
11. Eyal-Giladi, H. *Cell Diffn Dev.* **14**, 245–255 (1984).
12. Stern, C. D. *Development* **109**, 667–682 (1990).

* Workshop on Isotopic Anomalies, Clemson University, 11–14 November 1990.

present in the first-generation cloud induced to collapse by the AGB wind.

As both populations of grains expand in the bipolar outflow of the initial protostellar system and fall through the accretion shock of the second-generation star, the young grains would be unaffected whereas the older grains would be degassed. Of course, these grain populations would mix with the population of old gas-rich, interstellar SiC grains already present in the now collapsing cloud of the second-generation protostar. A natural consequence of this sequence is that all young, gas-poor grains found in primitive meteorites would have originated in the single AGB star which triggered the collapse of the first wave of protostars.

Although sequential star formation is known to occur for high-mass stars, and in fact is now occurring in the Orion Molecular Cloud (see Glenn White's News and Views article⁴) there is as yet no observational evidence that low-mass stars like the Sun form this way. In fact, I would suggest that there is no model for the origin of our own Sun which successfully predicts all of the complex isotopic and chemical relationships observed in the comets, meteorites and planets of our Solar System. Each new discovery, each new model adds another rich dimension to the increasingly complicated puzzle. It is even possible that the Solar System formed by some unlikely process which was significantly different from the usual low-mass star formation model. Unfortunately, it is difficult to assess the possibility that such an unusual event occurred without a thorough understanding of the archetypal process, and this has yet to be achieved.

The recent workshop on isotopic anomalies, alluded to above, was a first step on the road to a comprehensive understanding of the formation of low-mass stars such as our Sun. It brought experts on nucleosynthesis, meteoritics and star formation together with those familiar with the formation, destruction and properties of interstellar dust, facilitating a better understanding of the mechanisms by which circumstellar grains were transported into the Solar System and the processes which may have affected them along the way. Experts on interstellar dust were treated to close-up photos of individual pre-solar grains recently isolated from meteorites; and meteoriticists gained a new appreciation of the hostile nature of the interstellar environment in which such grains may have resided for several hundred million years.

Several presentations highlighted the potentially bewildering variety of isotopically anomalous elements isolated from meteorites, ranging in abundance from the more common elements such as hydrogen, carbon and oxygen, the isotopic variation of which can be directly

observed in stars, to extinct superheavy radionuclides whose existence is inferred from the presence of anomalous decay products. Each of these anomalies can tell us something about the pre-solar or nebular environment, yet each must be interpreted within the constraints imposed both by other isotopic systems and by our present understanding of grain formation, the interstellar medium and the process of star formation. It was clear from the meeting that we have a long way to go before the meaning and relevance of all of these constraints are understood. However, it was also evident that considerable progress has been made over the past decade if only by revealing how little about the process we really can be sure of.

I predict that over the next few years interdisciplinary studies related to the origins of planetary systems will yield a PALAEOLOGY

wealth of new observational discoveries. Many of these discoveries will challenge well established paradigms for the formation and early evolution of low-mass stars, planets, planetessimals and even dust. Such challenges represent true scientific progress. Cosmogony is beginning to leave its teenage years behind: we are no longer sure that we know how the Solar System formed, but are just beginning to realize how much there is which we still need to learn. □

Joe Nuth is at the NASA Goddard Space Flight Center, Greenbelt, Maryland 20770, USA.

1. Zinner, E., Amari, S., Anders, E. & Lewis, R. *Nature* **349**, 51–54 (1991).
2. Lewis, R., Amari, S. & Anders, E. *Nature* **348**, 293–298 (1990).
3. Gallino, R., Busso, M., Picchio, G. & Raiteri, C. *Nature* **348**, 298–302 (1990).
4. White, G. *Nature* **348**, 289–290 (1990).

New fossils and primate origins

R. D. Martin

THE known fossil record for the forest-living primates of the Eocene epoch (55–36 Myr ago) is relatively rich because of finds in the heavily worked Northern Hemisphere; because world temperatures were markedly higher, there was at that time a considerable poleward expansion of tropical and subtropical forests. Eocene primates are usually divided into two families — the first (Adapidae) contains numerous species collectively labelled 'lemuroids' because of a general resemblance to certain extant lemurs of Madagascar, the best-preserved fossils being of the North American genera *Notharctus* and *Smilodectes*; the second (Omomyidae) groups together several genera that have often been called 'tarsioids' because of

resemblances to modern tarsiers, but in this case the best-known forms, notably *Necrolemur*, are European. Hitherto, the partial skull of *Tetonius* was the only significant omomyid cranial material known from Eocene deposits of North America, and the only better preserved skull came from the later Oligocene form *Rooneyia*. The discovery of four skulls of the North American Eocene omomyid *Shoshonius*, reported by Beard *et al.* on page 64 of this issue¹, therefore dramatically augments our knowledge of cranial morphology in early 'tarsioids'.

It has been a common tendency² to link the Adapidae to the modern strepsirrhine primates (lemurs and lorises) and the Omomyidae to the modern haplorhine



How *Shoshonius* might have looked in the flesh, a reconstruction by Donna Braginetz.

Gathering dust

AEROGELS, the technological equivalent of meringue, comprise a rigid froth of walls encapsulating virtually nothing. The usual way of making them is to drive the solvent quickly out of a colloidal gel. But by pure chance, workers at the Karpov Institute of Chemical Physics in Moscow have found that aerogels can be grown directly from the vapour phase. A. A. Lushnikov, A. E. Negin and A. V. Pakhomov were observing the behaviour of tiny fractal aggregates (essentially specially prepared dust) suspended in air in a small vessel. To their surprise, as they describe it in *Chemical Physics Letters* (175, 138–142; 1990), after a few minutes the vessel became filled with a web of fine filaments. Even more surprisingly, the filaments started to grow in the interior of the vessel, not at its walls. Concluding in speculative style, the authors wonder whether such a process of aerial coagulation might be involved in stabilizing ball lightning.

Seals' meals

WRITING in the *Journal of the Marine Biological Association* (70, 829–840; 1990), G. J. Pierce and colleagues show that serological methods of faecal analysis, already applied in other ecological work, can be used to find out what seals eat in the wild; not the least of the advantages of the technique against an alternative method of examination of stomach contents is that there is no need to bother (or, at the extreme, kill) the animals under study. The authors particularly looked at the sandeel component of the seal diet, and compared the results from antisera raised against the remnants of sandeel protein in seal faeces to those obtained from analysis of sandeel bones in the faeces. They found that the two correlate well, and say that the approach is potentially applicable to other marine predators.

On the slide

THE experimental study of the avalanche instability of sandpiles has been carried further by J. Rajchenbach using a rotating cylindrical drum with transparent ends (*Phys. Rev. Lett.* 65, 2221–2224; 1990). The set-up is convenient because the transition between a succession of avalanches and continuous flow is a function of rotation speed; continuous flow supervenes at a critical slope θ_c . Rajchenbach confirms that the surface matter current in the continuous regime has the form $(\theta - \theta_c)^m$, and shows both by measurement and the extension of an argument due to R.A. Bagnold in 1954 that $m = 1/2$, whereas for a viscous liquid, $m = 1$. Studies such as these are now all the rage (see *Nature* 347, 225; 1990), not only because of their application in civil engineering but because of the belief that sandpiles are a model for a variety of self-organizing physical systems susceptible to instability.

primates (tarsiers, monkeys, apes and humans). There is a great deal of evidence to link modern tarsiers to monkeys, apes and humans³, and the resemblances between omomyids and tarsiers have often led to their inclusion in the haplorhines. These inferred relationships of adapids and omomyids have, however, been subject to criticism. Although a specific link between adapids and modern strepsirhines is supported by some new evidence^{4,5}, Gingerich's proposal⁶ that adapids — rather than omomyids — are linked to the origins of simians (monkeys, apes and humans) has recently received backing from Rasmussen and Simons^{7,8}.

At the same time, the supposed link between omomyids and modern tarsiers has been challenged, one of the main reasons being the demonstration that at least some omomyids, unlike modern tarsiers, possessed a gap between the upper incisors indicating that a rhinarium (a naked area of moist skin surrounding the nostrils) may still have been present^{9,10}. If that was the case, it is technically incorrect to refer to omomyids as haplorhines because they would have lacked one of the main derived features (suppression of the rhinarium) that links modern tarsiers to simians. Some authors^{10,11} accordingly believe that the omomyids branched away from the haplorhine line before the divergence between tarsiers and simians. Others^{9,12} have drawn the more radical conclusion that omomyids are related to strepsirhines rather than to haplorhines.

In several features, the new *Shoshonius* skull material confirms the picture already established for the Eocene Omomyidae. Relative to overall skull length, the orbits were large and indicate that *Shoshonius*, like *Necrolemur* and *Tetonius*, was nocturnal. A postorbital bar was present, as is typical for primates of modern aspect; but, as in other omomyid skulls, there was no development of a postorbital plate. The plate is present in modern tarsiers and simians, so this must be either a convergent development or a shared derived feature that developed after the omomyids branched away from the line leading to a common ancestor of tarsiers and simians. The snouts of the new skulls are unfortunately incomplete, so it is not yet known whether *Shoshonius*, like the European omomyids *Necrolemur* and *Pseudoloris*, had a gap between the upper incisors. Finally, in the ear region of *Shoshonius*, the ectotympanic ring was enclosed within the bulla and connected with its lateral wall by a set of perpendicular bony struts as in *Necrolemur* and *Rooneyia*.

The auditory bulla of *Shoshonius* also shows a number of apparently unique resemblances to modern tarsiers that might provide support for the idea that at least some of the Eocene omomyids were directly related to tarsiers¹. Like *Tarsius*,

Shoshonius had a suprategal foramen and a flange of the basioccipital overlapping the rear end of the medial bullar wall, and the posterior carotid foramen was in a ventrolateral position. As these resemblances to modern tarsiers are not found in other omomyids, the possibility arises that the family Omomyidae is in fact a polyphyletic group, containing some lineages that branched off earlier and at least one lineage (leading to *Shoshonius*) that is more closely related to *Tarsius*.

Although this evidence from the bulla is tentative, it certainly adds a new element to discussions of omomyid relationships. Interestingly, the snout of *Shoshonius* was even more reduced than in *Necrolemur* and *Rooneyia*, more closely approaching the condition in *Tarsius*. The relative size of the orbits was also greater in *Shoshonius* than in other known omomyids and closer to the extremely enlarged state found in modern tarsiers. Nevertheless, the new cranial evidence should ideally be backed up with evidence from the postcranial skeleton. The hindlimbs of modern tarsiers have a number of unusual features¹³, so a study of postcranial material from *Shoshonius* should be particularly revealing.

Should it emerge that *Shoshonius* is, indeed, specifically related to modern tarsiers, an interesting corollary emerges¹. The new skull material is about 50.5 Myr old and the divergence between tarsiers and simians would therefore have to be set at an even earlier date. Together with the recent discovery of the earliest known simian skull (*Catopithecus*), attributed to the late Eocene^{14,15}, this tends to confirm the suggestion³ that divergence times in the primate evolutionary tree may be markedly earlier than has been generally accepted. □

R. D. Martin is in the Anthropologisches Institut und Museum, Universität Zürich-Irchel, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland.

1. Beard, K.C., Krishtalka, L. & Stucky, R.K. *Nature* **349**, 64–67 (1991).
2. Szalay, F.S. & Delson, E. *Evolutionary History of the Primates* (Academic, New York, 1979).
3. Martin, R.D. *Primate Origins and Evolution* (Chapman & Hall, London, 1990).
4. Beard, K.C., Dagosto, M., Gebo, D.L. & Godinot, M. *Nature* **331**, 712–714 (1988).
5. Dagosto, M. *J. hum. Evol.* **17**, 35–56 (1988).
6. Gingerich, P.D. in *Evolutionary Biology of the New World Monkeys and Continental Drift* (eds Ciochon, R.L. & Chiarelli, A.B.) 123–138 (Plenum, New York, 1980).
7. Rasmussen, D.T. & Simons, E.L. *Folia primatol.* **51**, 182–208 (1989).
8. Simons, E.L. & Rasmussen, D.T. *Am. J. phys. Anthropol.* **79**, 1–23 (1989).
9. Schmid, P. thesis (Univ. Zürich, 1982).
10. Aiello, L.C. in *Major Topics in Primate and Human Evolution* (eds Wood, B.A., Martin, L.B. & Andrews, P.J.) 47–65 (Cambridge Univ. Press, 1986).
11. Cartmill, M. in *Evolutionary Biology of the New World Monkeys and Continental Drift* (eds Ciochon, R.L. & Chiarelli, A.B.) 243–274 (Plenum, New York, 1980).
12. Schmid, P. *Folia primatol.* **40**, 1–10 (1983).
13. Jouffroy, F.-K., Berge, C. & Niemitz, C. in *Biology of Tarsiers* (ed Niemitz, C.) 167–190 (Gustav Fischer, Stuttgart, 1986).
14. Simons, E.L. *Proc. natn. Acad. Sci. U.S.A.* **86**, 9956–9960 (1989).
15. Simons, E.L. *Science* **247**, 1567–1569 (1990).

At the eye of the storm

F. Michael Flasar

WHEN the two Voyager spacecraft flew by Saturn in 1980–81, they observed an atmosphere which, although banded like Jupiter's, had far less contrast and fewer conspicuous features. This is the usual situation; Saturn is fairly bland. Within the past two months, however, a giant 'storm' has erupted in Saturn's atmosphere. The phenomenon first appeared as a bright compact spot, just north of the equator, and was discovered on 25 September by S. Wilber, an amateur astronomer in Las Cruces, New Mexico. The word quickly got round, and Saturn became the object of close monitoring at several observatories.

According to R. F. Beebe of New Mexico State University and C. Barnet of the Goddard Space Center, astronomers who had been analysing images obtained with the University's telescope, the spot expanded in the ensuing days, maintaining an elliptical shape. By the beginning of October it extended 30° in longitude and 20° in latitude. At the same time a tail of more diffuse, cloudy material extended westwards from the spot, particularly from its northern edge. Several observers reported the appearance of a few localized brightenings within the vicinity of the spot. By mid-October the diffuse cloud extended around the planet within a latitude band extending from approximately 20° N to 9° S, and the original spot itself was much less distinct. Early in November, planetary astronomers were able to dedicate the Hubble Space Telescope for two days to observing Saturn. Despite the telescope's faulty primary mirror, clean images were possible, owing to Saturn's brightness and the consequent availability of digital enhancement techniques. The observable detail in the images (see figure) is much finer than that derivable from ground-based telescopes. One sees a distinct pattern of undulations along the northern edge of the cloudy band.

Some aspects of the storm admit a straightforward interpretation. At low latitudes the prevailing winds on Saturn are eastward. The original spot and the extended cloud are in a region of strong meridional shear. Time-lapsed tracking of discrete cloudy features in images obtained with the Voyager spacecraft (A. P. Ingersoll *et al.* in *Saturn* (eds T. Gehrels & M. S. Matthews) 195–238 (University of Arizona Press, 1984)) indicated that the eastward winds decrease from 500 m s⁻¹

near the equator to less than 200 m s⁻¹ at 20° N and S latitude. (Saturn has no visible surface, so the winds are pegged to its internal rotation rate, which is inferred from the observed periodic modulation in Saturn's kilometric radio emission, on the assumption that this is tied to the planetary magnetic field which rotates rigidly with the interior.) If the source of the cloudy material is near the equator, any

Saturn's Great White Spot, as seen by the Hubble Space Telescope.

meridional motion of the cloudy material would result in the material being stretched out around the planet in the matter of a few weeks.

Less obvious is why the storm occurs at all. Large tropical storms were also observed during Saturn's northern summer (the current season) in 1933 and 1876 (A. Sanchez-Lavega *Astr. Astrophys.* **185**, 315–326; 1987), or approximately every two saturnian years (Saturn's orbital period is 29.5 years). Similar storms occurred at northern midlatitudes during the same season in alternate saturnian years (1960 and 1903). The large storms have not been observed very often at other seasons, although obscuration of low latitudes by the rings could cause some bias. There may be a connection with the seasonally modulated solar heating, but the response of the atmosphere is complicated and not simply periodic or predictable. A. P. Ingersoll, an atmospheric dynamicist at California Institute of Technology, has, not entirely facetiously, termed the occurrences of these storms as "burps".

The typical bland appearance of Saturn is usually ascribed to the presence of condensate hazes, presumably of ammonia, that obscure the thick clouds below.

Jupiter, being warmer than Saturn, has less of this haze, and its cloudy features are more distinct. The eruption of the present saturnian storm suggests an instability, of as yet unknown origin, that has induced violent vertical motions, raising condensate to high altitudes of the atmosphere. Observations are currently being made at wavelengths in the electromagnetic spectrum other than the visible, and they may shed light on this process. Infrared observations, in particular, have the potential of determining the altitudes of the new clouds. They may also provide information on temperatures in the vicinity of the storm, which would be diagnostic of vertical motions.

Observers on the Hubble team are also measuring the meridional profile of the winds, by tracking cloud motions much as was done earlier with the Voyager images, to see if the storm has altered the prevailing wind pattern. The temporal variation of Saturn's winds is not well known. Historically, measurements of winds from ground-based images have indicated rough agreement with the winds inferred from Voyager. But because of Saturn's low contrast and the lower spatial resolution of these observations, the number of such measurements is small, and they typically have been based on the large clouds observable during the epochs of the storms. Understandably, they have large uncertainties.

Indeed, the driving force for the overall pattern of Saturn's winds is not understood, nor is it even known how deeply they extend below the visible clouds. (This state of affairs also holds for the circulation on the other outer planets.) The degree of steadiness inferred from the Hubble observations may shed some light on these questions.

Ideally, a giant saturnian storm would be subjected to observations at relatively close range by a battery of remote sensing instruments, such as those on the Cassini spacecraft, scheduled to begin orbiting Saturn at the end of 2002. We may not be so fortunate. The past great storms have dissipated in less than 100 days, and the current storm does not seem to be an exception. Saturn will be in its southern summer in 2002, and only one great storm has erupted during this season, at low southern latitudes in 1946. Hence, the long-term coverage afforded by ground-based and Earth-orbiting observatories may provide most of the clues to unravel this mystery. □

F. Michael Flasar is in the Laboratory for Extraterrestrial Physics, NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771, USA.

How do plants cope when they live in the shade?

Peter D. Moore

LIFE in the shade can be demanding for a plant, and assessments of gas exchange and carbon assimilation have often been used to compare the physiological responses to light intensity of plant species from sunny and from shady environments respectively. Gas-exchange studies by Ramos and Grace on the tree species of primary and secondary forests in Mexico, reported in *Functional Ecology*, now show that the species of primary forest largely behave in the manner of shade species, whereas those of secondary forests have 'sun plant' characteristics. This differentiation can be explained in terms of the establishment conditions of the two types of forest, species of mature, primary forests having to establish themselves under heavier canopies than those in disturbed or regenerating secondary forests.

Another, less easily explained, phenomenon has however been observed in the understorey herbs of both tropical and temperate forests, namely mottling of the leaves with spots or blotches. It is difficult to understand these features in terms of carbon balance. But, in another paper in the same journal, Givnish² proposes that they may have the advantage of camouflaging the plant, rendering it less liable to attack from herbivores.

Plants of shady habitats are often equipped with large, thin leaves that have lower maximal rates of carbon fixation, lower light compensation points, lower dark respiration rates, lower light saturation levels and steeper response curves to increases in light intensity at the low end of the scale than those of sun plants³. Ramos and Grace examined the gas exchange of four tropical forest species, one from primary forest, one from primary or late-stage secondary, and two from secondary forest. These plants, were grown in cabinets in which the light climate was adjusted to simulate shade and direct sunlight, both in terms of intensity and spectral composition. They were then tested for their carbon-fixation response to a range of flux densities, and it was found that the species of secondary forest have higher maximal rates of photosynthesis and are more sensitive to low light intensity than the primary forest species. These plants, which occupy habitats where open canopies and good light penetration prevail, can take advantage of the frequent, high-intensity sunflecks. A primary species must spend its juvenile existence under conditions of intense shade, however, and maximal growth rate is less important than is tolerance of such conditions.

Ultimately, what is important for the plant is whether the individual can maintain a positive long-term carbon balance in a given light environment, and it is difficult to translate experiments with leaf responses into an understanding of the balance for the entire plant. The allocation of reserves into different parts (leaf, stem, roots and so on) is of importance, as is the respiratory activity in these different organs. What one does not expect to find in a plant of shady environments is a feature that might be expected to reduce the efficiency of carbon fixation. This is precisely why the frequent association of variegated and blotched leaves⁴ with herbs of the woodland understorey is ecologically rather surprising. What possible advantage can a plant gain by giving up a significant portion of its photosynthetic surface? This is the question raised by Givnish in his paper.

Any additional spots of masking pigment, or the pale patches produced by a lack of chlorophyll, must reduce the photosynthetic efficiency of a leaf. Smith⁵ has proposed that the plant could be mimicking damage from leaf miners, which acts as a deterrent to further egg laying on the same leaf. But Givnish provides a wider interpretation of the feature, and claims that the leaf is gaining the advantage of crypsis in a microhabitat where the light environment itself may be patchy and complex in pattern.

Visual camouflage is known from the Plant Kingdom, perhaps most clearly in some succulent species, such as *Lithops*, the 'living stones' of desert areas. The avoidance of the attentions of herbivores could be of considerable benefit to a plant, but the cost of patterns of spots or variegation in terms of energy lost is a heavy one. One can appreciate that such markings may render a leaf less detectable; and they may also make it look smaller and thus less worth the effort of consumption, as in the case of the leaf patches on some clovers. But lost energy is only acceptable in evolutionary terms if the benefits of herbivore avoidance are adequate. We now need to know how effective and how important such avoidance would be in a forest-floor environment. □

Peter D. Moore is in the Division of Biosphere Sciences, King's College London, Campden Hill Road, London W8 7AH, UK.

1. Ramos, J. & Grace, J. *Functional Ecol.* **4**, 667–677 (1990).
2. Givnish, T. J. *Functional Ecol.* **4**, 463–474 (1990).
3. Boardman, N. K. A. *Rev. Plant Physiol.* **28**, 355–377 (1977).
4. Evanari, M. *Bot. Rev.* **55**, 106–139 (1989).
5. Smith, A. P. *Oecologia* **69**, 283–287 (1986).

Flesh pressure

AMONG the many aids to slimming, massage is one of the least credible, and yet most popular. Massage machines to pummel your adipose tissue, and local exercises designed to agitate it from beneath, are widely promoted. Daedalus wonders how this can make sense. Body fat, he observes, is in a constant state of metabolic turnover. As befits a reserve of chemical energy, it is slightly unstable. It is a mixture of glyceride esters, thermodynamically on the brink of hydrolysis to glycerol and fatty acids. Can mere mechanical battering push it over that brink?

Fat has a low density, and hydrolyses with a slight reduction in volume. So pressure must encourage its hydrolysis. Furthermore, says Daedalus, an adipose fat-droplet (like all small droplets) is potentially unstable. A small shrinkage must increase its internal pressure, driving it towards yet further shrinkage. So once destabilized by a brief burst of pressure, the fat-droplets in an adipose deposit should continue to shrink and hydrolyse faster and faster. Unlike conventional slimming, which soon runs into strong bodily resistance, pressure-treatment should trigger a wonderful, catastrophic, self-accelerating loss of fat.

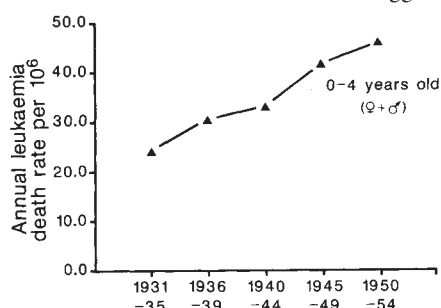
Conventional massage, however, could never achieve such a result. It must take at least ten or twenty atmospheres to trigger galloping fat-hydrolysis — a pressure inaccessible to the most violent masseur, unless he were also a deep-sea diver. Daedalus was on the point of inventing the submarine massage parlour, but soon realized that a compressed-air chamber would be more convenient. Working in such a chamber, DREADCO technicians are now squeezing corpulent volunteers into rapid and catastrophic collapse. Soon the DREADCO high-pressure slimming clinic should be open for customers.

This wonderful establishment will also offer active figure-shaping. The background pressure of the clinic itself will bring the customer's fat to the brink of dissolution; the exact regions to be tipped over that brink will then be selected by additional slight local pressure. Daedalus is designing a sort of 'iron maiden' figure-press, to squeeze those regions least desired by the customer. The fat at these points will be rapidly hydrolysed, while the products diffuse into the bloodstream. In unsqueezed regions, exposed only to the background chamber pressure, the products will tend to recombine into fat again. Thus women with the dreaded pear-shape syndrome will re-emerge as more fashionable hourglasses while men with central concentrations will redistribute them around their personal periphery. The dreams fostered by the slimming industry will come true at last.

David Jones

Leukaemia and wartime evacuation

SIR—Evans notes in News and Views¹ that Kinlen and colleagues² have provided evidence consistent with the hypothesis that childhood leukaemia may be a rare response to an unidentified infective agent. Kinlen and colleagues find that leukaemia deaths in 'rural' new towns are greater than those in 'overspill' new towns. Most at risk from the putative leukaemia infection resulting from this population mixing are the 0–4-year age group, which shows an observed/expected ratio for cases of 2.75 in the rural towns (compared with 0.95 for the overspill towns) in the early years following designation as new town. The authors suggest



that a higher population density and the greater diversity of places of origin in the rural (compared with overspill) towns would cause more severe leukaemia epidemics since there would be greater opportunity for 'herd mixing' infection but less herd immunity.

If it is true that there is an infective basis for childhood leukaemia than an earlier and more extreme population mixing would have contributed markedly to the national childhood leukaemia statistics. As a result of evacuation in Britain during the Second World War, 827,000 unaccompanied children, 524,000 mothers with children under 5 years old as well as 13,000 pregnant women migrated from major conurbations into rural areas between 1 and 4 September 1939 (ref. 3). In those 5 days alone, almost half of the nation's schoolchildren moved from the cities. Further migration and re-migration occurred during the course of the war.

In contrast, average population influxes into the 14 new towns studied by Kinlen and colleagues were a mere 20,000 (rural) and 38,000 (overspill) from 1951 to 1966. Yet the massive population mixing resulting from evacuation does not seem to have resulted in a surge of national childhood leukaemia⁴. The figure depicts data for leukaemia mortality of the relevant age group in England and Wales over the period 1931 to 1954. Against the unexplained steady rise in leukaemia mortality there is no evident sudden increase in the periods 1940–44 nor 1945–49.

Although extrapolation from global to local migration trends may be question-

able, the data suggest that herd mixing may be an inadequate explanation for local increases in leukaemia incidence. In another critique of the herd-mixing hypothesis (based on lack of specificity of increased childhood cancer) it was suggested that "environmental aspects of increasing urbanisation" could contribute to the increased incidence of childhood leukaemia, and other childhood cancers, in new towns⁵. I have proposed that locally higher exposure to benzene as a result of greater car usage encouraged by higher salary/cost-of-living ratios in the new towns is one such environmental factor⁶. The observation that residential traffic density in Denver correlates with child-

hood leukaemia (odds ratio = 4.7 for >10,000 vehicles per day and 2.1 for >500 per day)⁷ is also consistent with the hypothesis that childhood cancers may be linked to carcinogens associated with the internal combustion engine.

SIMON P. WOLFF

Department of Clinical Pharmacology,
University College and Middlesex
School of Medicine,
University College London,
London WC1E 6JJ, UK

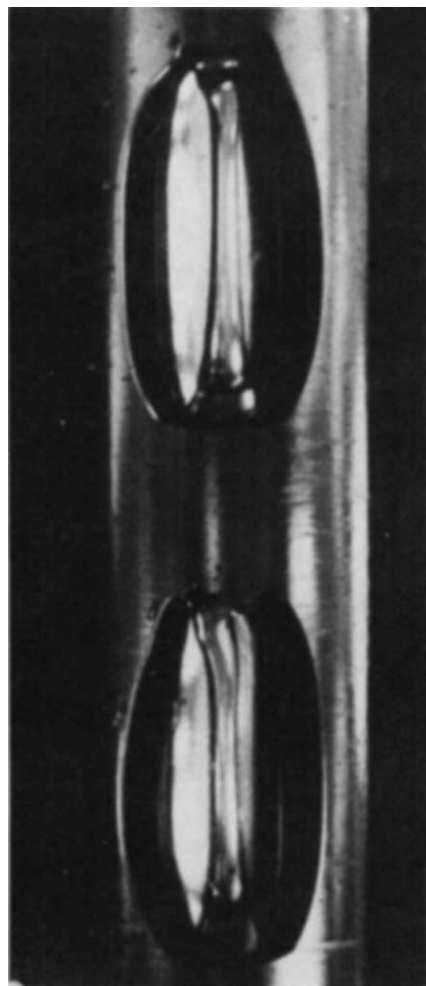
1. Evans, H.J. *Nature* **347**, 712 (1990).
2. Kinlen, L.J., Clarke, K. & Hudson, C. *Lancet* **336**, 577 (1990).
3. Minns, R. *The Domestic Front 1939–45* (Virago, London, 1980).
4. Court Brown, W.M. & Doll, R. *Brit. Med. J.* **1**, 981 (1961).
5. Muir, K.R., Graham, R., Parkes, S.E. & Mann, J.R. *Brit. Med. J.* **300**, 676–677 (1990).
6. Wolff, S.P. *Nature* **346**, 517 (1990).
7. Savitz, D. & Feingold, L. *Scand. J. Work Envir. Hlth* **15**, 360–363 (1989).

Toroidal bubbles

SIR—"Air-bubbles in water are usually rather deformed spheres. Daedalus reckons that toroidal bubbles should also be possible. He sees them as a form of two-phase vortex-ring — a toroidal air-bubble stabilized by the centrifugal force of the toroidally spinning liquid around it." This extract¹ from Daedalus's article 'Bubbling down' shows that amusing and speculative ways of thinking can be important in science. Toroidal bubbles really exist, and were observed by me in 1982 during investigations into two-phase downflow of air and very viscous mineral oil in vertical tubes. They are manifestations of a new flow pattern, for which we have proposed² the name 'stalactite flow'.

The photograph shows an example of toroidal bubbles inside a glass tube of internal diameter 25 mm for a liquid of viscosity 1.69 Pa s, surface tension 0.0363 N m⁻¹ and density 996 kg m⁻³. Toroidal bubbles have also been observed in a tube 15 mm in diameter. The flow pattern and the range over which it occurs are described in my thesis³ (which is in Polish); unfortunately, we have not been able to investigate this phenomenon in more detail because of the economic problems besetting science in Eastern Europe.

I also agree with Daedalus's opinion that his "interest is far from academic". The stalactite flow pattern could be applicable in multiphase flow reactors. Heat and mass transfer conditions should be favourable for toroidal bubbles, and may be better than in typical bubble column reactors. Stalactite flow could be especially useful for bioreactors which work with high viscosity liquids. In many biotechnology processes it is necessary to cultivate cells or bacteria in aerobic conditions with accurate control of



Toroidal bubbles in a tube of diameter 25 mm. temperature, and the cells and bacteria have to be protected against mechanical damage as much as possible. Multitubular reactors working under the stalactite flow regime could be a new step in developing bioreactors.

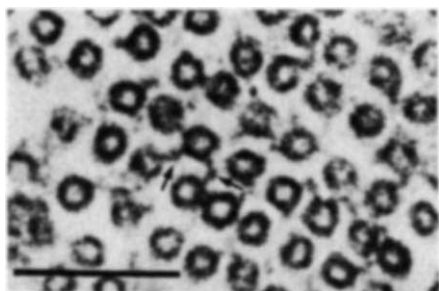
WOJCIECH SPISAK

Department of Heat Techniques and
Chemical Engineering,
Opole Technical University,
ul. ZSP 5, 45-233 Opole, Poland

1. Jones, D. *Nature* **343**, 122 (1990).
2. Troniewski, L. & Spisak, W. *Int. J. Multiphase Flow* **13**, 257–260 (1987).
3. Spisak, W. thesis, Wrocław Technical Univ. (1987).

Microtubule bundling in cells

SIR—Lewis and Cowan have amended¹ their hypothesis² that a small hydrophobic segment of neuron-specific microtubule-associated protein 2 (MAP2) is responsible for bundling microtubules in cells transfected with MAP2. The authors now suggest¹ that MAP2 is directly involved in bundling without using its hydrophobic carboxy terminus for this purpose. We suggest, instead, that the main effect of high-level expression of MAP2 in cultured fibroblasts is to stabilize dynamic microtubules, that the bundling that results is a direct consequence of this stabilization, and that bundling is mediated by an endogenous (as yet unidentified) bundling



Electron micrograph showing a cross section through a microtubule bundle in a fibroblast or presumptive myoblast treated with 10 μ M taxol for 3 days before fixation. Arrows, cross-bridges resembling those seen in cells transfected with MAP2 (see text). Scale bar, 0.1 μ m. (Reproduced from Fig. 15 of ref. 10 by permission of the Rockefeller Press; provided by Howard Holtzer and Camille DiLullo.)

protein, rather than by the exogenously introduced MAP2. There is evidence for this two-step bundling process of stabilization followed by bundling and against the idea that bundling is mediated directly by MAP2.

Three experimental treatments can cause bundling of microtubules in cultured cells. First, when fibroblasts overexpress transfected MAP2 or tau³ (neuronal-specific MAPs known to stabilize microtubules *in vitro* and *in vivo*), microtubule bundles are formed.

Second, taxol treatment causes almost immediate stabilization of microtubules *in vivo*⁴ and, with longer treatments, results in bundling in many cell types, including 3T3 and HeLa, the cell lines used in the studies by Lewis and colleagues⁵. Although at the light microscope level, some differences in the appearance and distribution of MT bundles may be observed between taxol-treated and MAP2- (or tau-) transfected cells, at the electron microscope level, one observes structural characteristics of the bundles that point to a common mode of bundling. Regardless of the origin of the bundles, the spacing between microtubules is 25 nm

(compare our figure with fig. 6c in ref. 2). Similar spacing within taxol-induced bundles has also been observed by other authors. Moreover, both MAP2- and taxol-induced bundles exhibit the same crossbridge structure; namely, a fine filament with a central thickening. The similarity in microtubule spacing and crossbridge structure in MAP2- and taxol-induced bundles strongly suggests a common mode of bundling independent of MAP2, as MAP2 is absent from cells exhibiting taxol-induced bundles. Taxol-treated 3T3 cells containing microtubule bundles do not contain MAP2 or tau (G. G. G., unpublished results).

Third, microinjection of a nonhydrolysable GTP analogue⁶, which stabilizes microtubules *in vitro*⁷, into cells containing neither tau nor MAP2 induces microtubule bundles, further evidence for a bundling mechanism dependent on microtubule stabilization but independent of MAP2 and tau.

Thus, three independent treatments result in microtubule stabilization and bundling in cultured cells. Additional evidence against the hypothesis that MAP2 or tau directly mediates bundling is that both proteins behave as monomeric species in solution (ref. 8; J. C. B., unpublished data). The identity of the endogenous factor that is responsible for bundling is unknown, although dynamin, a recently identified bundling protein⁹, is a possible candidate. Stabilization of microtubules may stimulate bundling by increasing the longevity or concentration of microtubules, the abundance of non-centrosomal microtubules or biosynthesis of tubulin.

Given the relationship between stabilization and bundle formation, our interpretation of the deletion analysis of MAP2 described by Lewis and Cowan¹ is that those MAP2 constructs that did not induce bundling failed to do so because the transfected MAP2 did not stabilize microtubules adequately to allow endogenous bundling factors to act. MAP2 was co-localized with microtubules in some of these instances, in these cases we would expect that MAP2 stabilizes microtubules only slightly. Alterations in MAP2 synthesis, steady-state level or binding affinity could explain the less dramatic effects of these constructs on the microtubule cytoskeleton. Thus, Lewis and Cowan have not demonstrated a

direct function for MAP2 in bundling, but they have raised the interesting question of the role of stabilization in generating microtubule bundles.

STEVEN J. CHAPIN

JEANNETTE CHLOE BULINSKI

GREGG G. GUNDERSEN

Department of Anatomy & Cell Biology,
Columbia University,
College of Physicians and Surgeons,
New York,
New York 10032, USA

Name dropping

SIR—Dryja in News and Views¹ discussed several examples of inherited retinal degradation caused by mutations in the genes encoding known photoreceptor-specific proteins. He said in passing in his article that the gene that is mutant in the *rd*s (retinal degeneration slow) mouse encodes the rod disk membrane protein peripherin. This statement misleadingly suggests that peripherin is a well-characterized retinal protein.

The *rd*s gene was originally cloned by myself and co-workers² by using a 'reverse genetics' approach, and we determined the sequence of the wild-type *rd*s messenger RNA. A sequence database search showed that the *rd*s protein was novel. We next showed biochemically that the *rd*s product is a glycoprotein associated with disk membranes whose distribution is confined to photoreceptor outer segments³. Molday *et al.* raised monoclonal antibodies against rod outer-segment disks and observed an antigen of unknown identity in the disk margins⁴. They named this protein 'peripherin' because of this distribution. They later showed⁵ that this antigen is the bovine homologue of the *rd*s protein.

To what extent does giving a name to an otherwise unknown protein based on a relatively superficial characteristic, such as its position within an organelle, amount to a functional characterization? I believe that the 'claiming' of novel proteins of unknown function by naming them adds little to our understanding. I suggest that we leave 'peripherin' for the unrelated neurofilament protein that already has that name, particularly as our work³ on the *rd*s protein does not support its distribution as confined to the disk periphery.

GABRIEL H. TRAVIS

Department of Psychiatry,
University of Texas Southwestern
Medical Center,
Dallas, Texas, 75235, USA

1. Lewis, S.A. & Cowan, N.J. *Nature* **345**, 674 (1990).
2. Lewis, S.A. *et al.* *Nature* **342**, 498–505 (1989).
3. Kanai, Y. *et al.* *J. Cell Biol.* **109**, 1173–1184 (1989).
4. Gundersen, G.G. *et al.* *J. Cell Biol.* **105**, 251–264 (1987).
5. Piperno, G. *et al.* *J. Cell Biol.* **104**, 289–302 (1987).
6. Wehland, J. & Sandoval, I.P. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1938–1941 (1983).
7. Sandoval, I.V. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **74**, 4881–4885 (1977).
8. Cleveland, D.W. *et al.* *J. molec. Biol.* **116**, 227–247 (1977).
9. Shpetner, H.S. & Vallee, R.B. *Cell* **59**, 421–432 (1989).
10. Antin, P.B. *et al.* *J. Cell Biol.* **90**, 300–308 (1981).

1. Dryja, T.P. *Nature* **347**, 614 (1990).
2. Travis, G.H., Brennan, M.B., Danielson, P.E., Kozak, C.A. & Sutcliffe, J.G. *Nature* **338**, 70–73 (1989).
3. Travis, G.H., Sutcliffe, J.G. & Bok, D. *Neuron* **6**, 1–10 (1991).
4. Molday, R.S., Hicks, D. & Molday, L. *Invest. Ophthalm. Vis. Sci.* **28**, 50–61 (1987).
5. Connell, G.J. & Molday, R.S. *Biochemistry* **29**, 4691–4698 (1990).

Ancient incredulity

Henry Gee

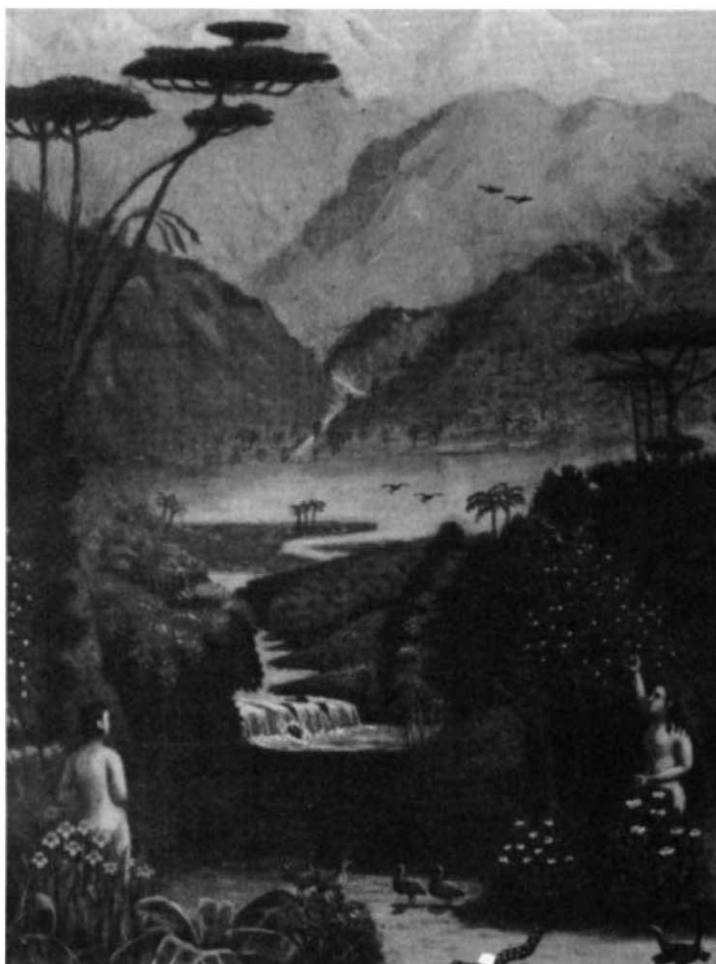
The Journey from Eden: The Peopling of Our World. By Brian Fagan. *Thames and Hudson*: 1990. Pp.288. £12.95, \$22.50.

OF the many great institutions Britain has to offer, one of the finest is the Caption Competition in the humorous weekly magazine *Punch*. Readers are invited to put new words to an old cartoon, sometimes one that was printed in *Punch* more than a century before. The winning entries are printed each week, and the contributors receive, in the best traditions of the British amateur, a small honorarium. In one, a fat victorian farmer criticizes a recently commissioned portrait of his prize bull. For the same money, he says, Picasso paints two heads.

These old cartoons are littered with the impedimenta of a bygone age: anti-macassars and horse-buses, liveried servants and bedwarmers, cigar-cutters and smoking-jackets. We can relate to them because the modern world is still enriched with victorian memories and ideals. But what would we know about the victorians if *Punch* cartoons represented the entire record of their era? This is exactly the problem faced by palaeolithic archaeologists, who try to reconstruct — from precious little material evidence — the lives and times of people thousands of times as remote from our age as are the victorians. As such, *The Journey from Eden*, an archaeologist's honest assessment of such evidence as we have for the mode of origin and subsequent global spread of anatomically modern humans, represents the credulous in pursuit of the tenuous.

Brian Fagan sets out to discuss the two main arguments in today's Genesis debate. Most researchers agree that *Homo erectus* spread from Africa into Eurasia around 700,000 years ago, evolving into a number of derived forms similar to modern human beings, and which Fagan calls archaic *Homo sapiens*. Some researchers think that modern human populations essentially reflect descent from these regional isolates, advocating the so-called 'Candelabra' hypothesis. Other researchers, however, espouse the

'Noah's Ark' or 'Garden of Eden' idea in which all anatomically modern humans (what Fagan calls *Homo sapiens sapiens*) derive from one or a small number of sub-Saharan populations that spread from an African 'Eden' about 100,000 years ago, replacing the archaic forms.



Primaevial Eden — was the true Eden the African savanna?

The crew of 'Noah's Ark' is currently in the ascendant, thanks to two recent and well-publicized pieces of evidence. First, thermoluminescence dates of stone tools associated with modern human skeletons found in Israel show them to have lived in the area over 90,000 years ago (40,000 years earlier than near-eastern Neanderthals) and just in time to have been the first modern humans to have left Africa.

Second, an analysis of maternally inherited mitochondrial DNA (mtDNA) in a small sample of modern humans suggests that all modern human mtDNA can be traced to a single ancestral mitochondrial genome originating in Africa

somewhere between 100,000 and 200,000 years ago.

Fagan refers to these and other recent developments, and in a brisk canter round the world, reviews the major archaeological evidence as well. His own position is never made explicit, but he seems to favour a mixture of both Noah's Ark and Candelabra, with a definite bias to the former. But much of his arguments seem curiously muddled, so much so that one might, on reading the book, be forgiven for thinking that scientific archaeology is a contradiction in terms.

The treatment of the mtDNA evidence is a case in point. Fagan seems to accept these data without question as evidence in favour of the Noah's Ark hypothesis (but to be fair, many other people have as well).

Picture is taken from *The Journey from Eden*

The argument is based on a diagram in a now-classic *Nature* paper (Cann *et al.*, 325, 31–36, 1987). This is simply a dendrogram summarizing the similarities among restriction endonuclease digests of 133 samples of modern human mtDNA, broken down by region of origin. It suggests that (most) African samples form a cluster of heterogeneous mtDNAs, in contrast with the more homogeneous mtDNAs from the rest of the world.

Because the authors drew the diagram as a rooted dendrogram to emphasize this distinction, the world at large took it literally, as a family 'tree'. After this misreading, it is easy to see that the Eurasian mtDNA and the African mtDNA types bifurcate quite near the root. This node has been interpreted as the point at which modern humans left Africa. Even if true, this was just as likely to have been the result of random population

interchange as wholesale migration. The fact that such a population would have trodden a route that presaged the Biblical Exodus has a mythic significance that has been ignored.

Ignored, too, is the fact that the extent of sequence divergence over the entire tree is less than 0.6 per cent, leaving room for error. Is the African-Eurasian difference by which people have set such store threatened by such error? What is the likelihood of alternative 'trees'? I have no idea, and by the uncritical tone of his book, neither has Fagan.

Because mtDNA is maternally inherited, commentators reasoned that

somewhere in Africa 100,000 years ago was a woman, a primaevae Earth Mother and the ancestor of us all. Eve (who else?) became the biggest media superstar who never was. A 1980s woman of the Stone Age, she raised the children and supported a career gathering eco-friendly vegetarian wholefoods, and tolerated men for as long as it took to assure the future of her own mitochondria, but not a moment longer. Adam, one notes, never gets a mention. Fagan, to his credit, balks at the notion of a single Eve figure who is identifiable as such. "In the sense that there was once a single person one could point to as the identifiable, single ancestor of *Homo sapiens*, the answer is no" he says (page 32). But a few sentences on, he notes that "there may, indeed, once have been a single Eve, but we can never hope to identify her in person".

Fagan's taxonomy, too, is woolly. The formal differences between *Homo erectus*, *Homo sapiens* and *Homo sapiens sapiens* are never made clear, apart from a note that the line between *H. erectus* and *H. sapiens* is hard to draw. One might have thought that rigorous definitions of all taxa concerned would have been essential. It could be that there are no agreed answers to this, in which case Fagan should have been much more explicit about the taxonomic problems posed by traditional archaeological methods.

Traditional archaeology, though, finds Fagan on home turf. His discussion on toolmaking among early *H. sapiens* in Africa was absorbing, partly because it seemed so new. *Australopithecus*, *H. habilis* and the Leakeys have cast such long shadows over African palaeoanthropology that it was refreshing to be reminded of sites such as Broken Hill and Klasies River Mouth. Identifying workmen from their tools, though, is rather fraught, a bit like divining the breed of a cat from its smile. The only taxonomy here is cultural, and this is dangerous. Referring to the early colonization of Australasia by boat (page 137), Fagan says that "the expertise to cross water out of sight of land, however accidental initially, was a distinctive skill of *Homo sapiens sapiens*". Maybe, but only if we find one buried with a map and compass.

The Journey from Eden is hobbled by a casual attitude too reticent to question assumptions that have a central bearing on the book's entire theme. Frequent lapses of consistency and logic — not to mention style — suggest a book written in a hurry. To document the spread of modern humanity is a worthy aim, and it deserves a far more measured and critical appraisal. Clear thinking and attention to detail are two Victorian values worth keeping: for the same money, Picasso would have given you two heads. □

Henry Gee is on the editorial staff of *Nature*.

Good time guide

Michael McElhinny

A Geologic Time Scale 1989. By W. B. Harland, R. L. Armstrong, A. V. Cox, L. E. Craig, A. G. Smith and D. G. Smith. Cambridge University Press: 1990. Pp. 263. Hbk £25, \$49.50; pbk £11.95, \$19.95.

IN the earth sciences the study of time is one of the most fundamental endeavours necessary for the development of almost all aspects of the science. There have been many successive attempts by various authors to produce definitive time scales covering the whole of Earth history, perhaps the most notable of these being *A Geologic Time Scale 1982* by Harland and coauthors, the predecessor of this book. The new and updated version is quite outstanding and represents a model of scientific synthesis and presentation.

The work develops and assesses a new calibration of the geological time scale using a new database up to 1988. It adopts the same style and uses and develops similar methods to the previous volume, but it has been entirely reworked. It should be noted that the 1982 *Time Scale* was based mostly on data compiled before 1976. The 1989 *Time Scale*, however, not only presents the state of the art in 1989 but the data assembled provide a source of reference which will serve for some years. The list of nearly 1,000 cited and other references provides an excellent bibliography.

Some of the most important and significant aspects of this volume relate to the way in which the database of isotopic determinations has been refined by tests both stratigraphic and geochemical. Many of the age determinations used for the 1982 *Time Scale* have been rejected and the database now includes Cenozoic data (previously the Hardenbol and Berggren Cenozoic scale of 1978 was simply incorporated unchanged). About 700 isotopic determinations are listed. The chronostratigraphic scale is calibrated applying and developing the chronogram method introduced with the 1982 *Time Scale* and based on the method used to estimate the ages of magnetic reversals from radiometric data. The process resulted in 127 chronograms of which 125 are reproduced in an appendix. A major advantage of this method is that a new scale can be produced automatically in a few hours incorporating new data and varying the data input. This allows comparison of the effects of the inclusion and exclusion of different classes of data. Finally the magnetostratigraphic scale has been updated and this in turn enables some refinement to the resulting time scale.

By showing the method of construction the reader can assess how the new time scale may be modified as more critical

data become available. The estimated uncertainties in the time scale values adopted are significantly less than in previously published estimates. Changes from the 1982 *Time Scale* are most significant in the Jurassic-Triassic and during the Upper Carboniferous. The base of the Cambrian has been reduced from 590 to 570 Myr BP and seems likely to be reduced further in the future. The Vendian has been given substantial treatment, resulting in a reduction of the estimated age of its base from 670 to 610 Myr BP.

This work will be of fundamental importance to almost every earth scientist and I highly recommend it to all both as compulsory reading and as an essential part of their personal library. A newly designed coloured wall chart of the 1989 *Time Scale* is also available as a separate publication and is the sort of thing one could expect to see on the wall of everyone's office. □

Michael W. McElhinny, Gondwana Consultants Pty Limited, 112 Sealand Road, Fishing Point, New South Wales 2283, Australia.

Critical errors

Ted Forgan

Fundamentals of Superconductivity. By V. Z. Kresin and S. A. Wolf. Plenum: 1990. Pp. 231. \$42.50.

THERE must be many readers, intrigued by the phenomenon of high-transition-temperature (high- T_c) superconductivity, who would be interested in a book which introduces superconductivity and explains the significance of the new high- T_c materials. With the title *Fundamentals of Superconductivity*, and the claim in the preface that it is aimed at a broad audience, Kresin and Wolf's book might appear to fulfill this need.

The authors state that they have concentrated on the qualitative aspects, so that the reader will not get bogged down in details. To me, many of their explanations seem not so much qualitative as slipshod, particularly in the theory chapters. For instance, to demonstrate the existence of the superconducting energy gap, the authors quote formulae for the low-temperature heat capacity, but do so on three separate occasions, and each time give a different formula. The reader is left unsure whether only two or all three of these formulae are incorrect. There are also startling changes in the difficulty and degree of explanation: advanced topics appear without warning next to laboured treatments of basic ideas.

On too many occasions, confusion is caused by misprints in text and equations, and a few of the diagrams are rendered almost incomprehensible by inadequate

or misplaced captions. There are occasional glaring errors, for example, on page 40 in a purported proof of quantization.

Once past the introductory theory, the book becomes more readable, with useful surveys of the occurrence and mechanisms of superfluidity from organic superconductors to neutron stars. There are other chapters on thin-film techniques, and on how one makes reliable measurements of superconducting properties such as upper critical field and critical current density. I was surprised not to find a proper discussion of granular effects in the latter context.

Another chapter gives a rapid survey of applications of superconductivity; it is probably too brief to give a clear indication of the status of many of the applications. Once again there are items which are plainly wrong, such as the description of how a flux transformer works in the input circuit for a superconducting quantum interference device.

The book concludes with an account of high- T_c materials and an outline of a possible theory to explain their properties. We are promised an updating in a future edition: my impression is that other parts of the book are in more urgent need of revision. I learnt some things from reading this book, but I would not recommend new readers to begin with it. □

Ted Forgan is in the School of Physics and Space Research, University of Birmingham, PO Box 363, Birmingham B15 2TT, UK.

What the brain computes

Sara J. Shettleworth

The Organization of Learning. By Charles R. Gallistel. MIT Press: 1990. Pp.648. £40.50, \$45.

In the past 20 years or so the study of learning in animals has moved from almost exclusive emphasis on simple associations in pavlovian and instrumental conditioning to consideration of apparently more cognitive processes like those required for spatial orientation and timing. Theoretical developments have not kept up. The conventional textbook treatment of learning still starts with habituation, moves on to simple associative learning, and concludes with 'other approaches' — studies of animal cognition on the one hand and analyses of what and how animals learn in their ecological niches on the other. How these should be integrated theoretically with accounts of simpler learning is seldom clear, although connectionist models can be taken as an attempt to show how an associationist integration might be possible. Attempts to

unravel the neural basis of learning show a similar discontinuity between cellular studies of habituation and simple associations on the one hand and, on the other, studies of how such complex aspects of the world as the structure of visual and auditory space are represented in animal brains. In *The Organization of Learning* Gallistel attempts to turn the study of learning on its head by discarding associationism and reorganizing behavioural and neural analyses of learning around a "computational representational approach". Whether he has completely succeeded only time will tell, but in the meanwhile this book will be difficult to ignore.

Here is a book about learning that starts with animals functioning in their worlds: hummingbirds foraging for nectar, desert ants finding their way home. Complex behaviours such as these must, Gallistel insists, be guided by detailed representations of the relevant aspects of the environment and neural computations on them. Representation here has a notably richer and more explicit meaning than in most studies of animal cognition where, for example, asking how an animal represents a reinforcer or a to-be-remembered signal means little more than discovering what aspect(s) of the event are encoded. In Gallistel's terms, "standing for" an event is a maximally impoverished sort of representation. What interests him are functioning isomorphisms between neural events (as evidenced primarily in behaviour) and aspects of the world. Perhaps the most familiar example is the hypothetical cognitive map, which allows an animal successfully to take new routes between familiar places. Accordingly, nearly a third of the book is devoted to spatial representations.

The starting place is a consideration of the computations that animal navigators implicitly must perform, in the form of a fascinating exposition of human navigational systems. Here, however, is a danger in Gallistel's approach. If an animal performs a feat that a human navigator would achieve by the use of specifiable computations on observed data, does it follow in any useful sense that the animal's nervous system must be performing the same com-



Sex change — the red anthia is a tiny fish which lives and feeds in the coral fringes of the Red Sea. Females outnumber males by about 100 to 1, and when a male dies one of the largest females reverses its sex and becomes a male. The picture is taken from *Arabia: Sand, Sea and Sky* in which Michael McKinnon describes the various changes that have shaped Arabia and its marvellous wildlife in recent millennia. Published by BBC Books, price is £15. □

putations?

That it may be misleading to assume an answer in the affirmative is well illustrated by the ongoing history of research on foraging behaviour. Optimal-foraging models can be used to compute how animals ought to respond to variations in feeding rate, but there are now numerous demonstrations that close-to-optimal behaviour can sometimes be achieved by simple rules of thumb nothing like the theorist's computation. For example, computation shows that as food becomes sparse, animals should begin to take prey which they would reject in better times. A changing rate of prey encounter can be tracked by an animal with a threshold for accepting prey which falls with time since the last prey encountered. But an animal that tracks local encounter rates in this way is computing numbers of items per unit time only in a degenerate and representationally uninteresting sense. Such data notwithstanding, Gallistel uses a chapter on rate to suggest that animals use rich representations of time and number, functioning isomorphisms of these quantities in the world, to generate (by division) representations of rate.

Not only has Gallistel taken on timing,

counting and instrumental choice behaviour, he proposes — among other things — a new approach to simple associative learning based on computation of rate as well as a scheme for representing vector quantities in the nervous system. People working in some of the areas which he reinterprets, like pavlovian conditioning, may be irritated by his selective treatment of the literature (that the book is as long as it is reflects the extensiveness with which a few telling examples are discussed rather than any attempt at a comprehensive review) or baffled by the unfamiliar mathematics. Nevertheless, the material in this book has to be reckoned with by anyone involved in the behavioural or neural analysis of learning. It is a unique and exciting attempt at a unified account of how animals store and use information about their worlds. □

Sara J. Shettleworth is in the Department of Psychology, University of Toronto, Toronto, Ontario M5S 1A1, Canada.

Ill-conceived project

Richard Ruquist

Nuclear Dynamite: The Peaceful Nuclear Explosions Fiasco. By Trevor Findlay. *Brassey's Australia*: 1990. Pp.339. £19.50, \$35.

In the 1960s the US Atomic Energy Commission (AEC) was seemingly more concerned with the survival and maturation of project Plowshare, a programme to develop peaceful nuclear explosions (PNEs), than it was with the proliferation of nuclear weapons. Non-proliferation was the concern of the US Arms Control Disarmament Control Agency (ACDA). In ACDA's view "... the technology of making nuclear explosive devices for peaceful purposes is indistinguishable from the technology for making nuclear weapons. ..." The consequential schizophrenic US policy, fostered by indifferent presidents and a split bureaucracy that pit the AEC against the ACDA and the Bureau of the Budget, interfered with the US acceptance of important arms control treaties, namely the Limited Test-Ban Treaty and the Non-Proliferation Treaty, and it helped to subvert a comprehensive test-ban, according to Trevor Findlay in *Nuclear Dynamite*.

Plowshare, like other US nuclear programmes, was born and nurtured by a scientific, bureaucratic and political alliance. The principal personalities of this alliance were: Edward Teller of the University of California Lawrence Livermore Laboratory, famed inventor of the fusion bomb, who helped initiate the Strategic

Defense Initiative (SDI) programme; Glenn Seaborg, the Director of the AEC from 1961 to 1971, who shared the Nobel Prize for the discovery of plutonium; and the Californian congressional members of the Joint Commission of Atomic Energy, Republican Craig Hosmer and Democrat Chet Holifield, its long-standing chairman. Findlay credits Teller with the longevity of this "ill-conceived project". No doubt a relatively small budget that peaked at \$18 million in 1968 also helped, not counting untold amounts borrowed from the weapons programme. The alliance was backed by a strong constituency in science and industry. In the end, both became disenchanted. But not Teller. He seems to hold true to a desire expressed to Enrico Fermi in 1945 — to develop the fusion bomb to extend our power over natural phenomena.

The history of Plowshare is complete. Trevor Findlay presents it from beginning to end in 16 chapters, five appendices and a wealth of footnotes. Included with this history are discussions of similar programmes in Australia, Argentina, Brazil, India and the Soviet Union, all stimulated by over zealous promotion of Plowshare. One chapter is devoted to the Soviet programme, a programme that dwarfed that of the United States in all aspects, yet was only initiated in 1964 in response to Plowshare, following a long period of Soviet anti-PNE sentiment. Besides numerous nuclear test-site experiments, the Soviet programme has consisted of at least 110 field-tests, whereas the US programme had only 35 tests in total, and three ill-fated field-tests.

This is not a dispassionate history, for Findlay speaks from the perspective of an arms controller. A former member of the Australian disarmament delegation from 1979 to 1984 and currently at the Australian National University Peace Research Centre, he makes the case against nuclear dynamite in fascinating detail. With a light and sometimes cynical touch, he usually lets the history of Plowshare speak for itself. He discusses various applications of PNEs and all of the major US experiments in detail. But *Nuclear Dynamite* is mainly about the politics of Plowshare: how it was initially used to argue, to no avail, against nuclear test moratoriums and treaties; how President Johnson took political advantage from the programme; and how the Limited Test-Ban Treaty continually hemmed in the programme.

The possibility of building a new isthmus sea-level canal by nuclear excavation was used by President Johnson to keep Panama at bay while renegotiating a new Panama Canal Treaty. This was the life-blood of Plowshare in the late 1960s. As part of this scene, Findlay presents an intriguing account of a Teller-initiated Australian harbour-creation project, a trial run for the Panama Canal project,

that failed to materialize for a myriad of technical and economic reasons.

The thrust of the Plowshare programme appeared to reach fruition in 1970, when a provision that institutionalized PNEs was incorporated in the Non-Proliferation Treaty. But the programme then immediately encountered a series of setbacks. In the summer of 1970 the first thorough study of nuclear excavation was released, the [Panama] Canal Commission Report. It demolished the technical and economic justification for surface applications. Actually the programme had already gone underground following the 1968 35KT Schooner cratering experiment in which significant levels of radioactivity were released into the atmosphere, thereby violating the Limited Test-Ban Treaty. The setbacks continued as a series of underground tests were cancelled by feasibility studies or by opposition from citizens. Three field experiments were performed to stimulate the enhanced release of natural gas, but these provoked controversy, proved ineffective, and the gas industry lost interest. A project to produce underground storage facilities was killed by citizen unrest.

Findlay admits to one possibly useful application of PNEs, that of building deep underground storage cavities for radioactive waste. Ironically, this concept was killed by the AEC itself because it was contrary to established policy. All testing ended in 1973. Congress abolished the AEC in 1974 and forbade the funding of any field testing of nuclear explosives for the recovery of gas or oil.

The ground-swell of citizen concern for the environment was beginning to take effect. Finally, the 1974 test of a nuclear device in India, for the expressed purposes of peaceful nuclear explosions, fixed world opinion against PNEs. Findlay credits the supersession of the arms controllers to grass-roots popular support. A weakness of this book is that it pays very little attention to the efforts of various arms control and environmental organizations, both public and private, to develop that support. Findlay does make a strong case for a comprehensive test-ban that abolishes PNEs, rather than a separate PNE treaty that might make them legitimate. He believes that the time for a comprehensive treaty has arrived.

After all, the Soviets in 1977 were prepared to accept a suspension of all underground tests including a moratorium on PNEs. Since then the Soviet excavation programme has been shelved by the direct intervention of Gorbachev. Unfortunately, the Reagan administration decided not to resume negotiations in 1982 in response to the needs of the nuclear weapons industry. □

Richard Ruquist, 27 Edgewood Road, Lexington, Massachusetts 02173, USA.

Localized sex in bacteria

John Maynard Smith, Christopher G. Dowson & Brian G. Spratt

Electrophoretic studies suggest that bacterial populations consist of a number of independent clones and that genetic recombination is rare in nature. But DNA sequencing reveals that individual bacterial genes have a mosaic structure that could have arisen only by recombination. How can these observations be reconciled and what is their relevance to prokaryotic evolution?

IN 1928, Griffith¹ showed that, if killed virulent and living non-virulent pneumococci were injected into a mouse, living virulent bacteria could be recovered. This discovery of bacterial transformation was the first step on the road to molecular biology, but, until recently, little was known about the role of transformation and other mechanisms for the exchange of chromosomal genes in evolution. Here we contrast recent evidence that recombination between gene loci is rare in bacteria, so that bacterial populations consist of a series of independently evolving clones^{2,3}, with studies that demonstrate horizontal transfer of chromosomal genes in bacteria, and discuss the relevance of these observations to prokaryotic evolution.

Clonal structure of bacterial populations

Simultaneous electrophoresis of a number of enzymes in natural isolates of *Escherichia coli*^{4,5} has shown that this species is highly variable: 94% of loci were found to be polymorphic, with a mean diversity per locus of 0.34–0.54. Given this diversity, and the number of loci studied, an immense number of distinct genotypes might be distinguished, but in practice only a small fraction are found. Some of these electrophoretic types are found repeatedly, in different continents, over many years. Estimates of the number of clones present vary from 100 to 1,000. Recombination affecting large regions of chromosome, mediated by chromosome transfer during conjugation, must be too rare to break up these widely distributed clones.

A similar clonal population structure exists in most other bacterial species, including the naturally transformable *Haemophilus influenzae*⁶ and *Neisseria meningitidis*^{7,8}, where recombination might be expected to occur more frequently. For example⁷, 15 enzymes in 688 isolates of *N. meningitidis* were polymorphic, with an average of seven electrophoretic alleles per locus. Yet 61 electrophoretic types were found more than once, and 19 were isolated from two or more countries, and up to 15 years apart. The stability of particular clones is striking: only seven of them, each consisting of a cluster of closely related electrophoretic types, caused many of the epidemics of serogroup A meningococcal disease over the past 70 years⁸. Outbreaks of serogroup B meningococcal disease in Europe, the United States, Cuba and Chile during the past 15 years were caused by members of a single clone⁹.

Mosaic structure and localized recombination

The DNA sequences of genes from closely related bacteria show a mosaic structure, with regions of high similarity interspersed with regions of up to 20% nucleotide difference, indicating that local recombination events involving a few hundred base pairs must have occurred. This is well shown in the evolution of penicillin resistance in *Streptococcus pneumoniae* (Griffith's pneumococcus).

Penicillin kills bacteria by inhibiting enzymes involved in peptidoglycan synthesis—the penicillin-binding proteins or PBPs¹⁰. The PBPs of penicillin-resistant pneumococci have decreased affinity for penicillin¹¹. The gene encoding one of these enzymes, PBP2B, has been sequenced from 6 sensitive and 14 resistant strains¹². The genes from the sensitive strains,

collected over a period of 50 years from three continents, were very similar: the greatest difference between two strains was 14/1,453 nucleotides, causing two amino-acid differences. An unlinked gene, that for amylomaltase, was sequenced from two sensitive and six resistant strains: only 3/650 nucleotide sites were polymorphic. Thus *S. pneumoniae* is genetically rather uniform.

Figure 1 shows the structure of the PBP2B genes of the penicillin-resistant strains. They fall into classes A and B. The class B genes have been found in almost all the resistant-serotype 23 strains that predominate in Spain and the United Kingdom and the sequences of the genes from five of these strains, isolated between 1984 and 1988, were identical. The genes differ from the gene of the sensitive strain R6 by 57/274 (20.8%) of nucleotides in a central region (codons 399–490), but by only 2.1% in the rest of the gene (Fig. 1*h*).

The nine class A genes were found in penicillin-resistant pneumococci isolated from various countries between 1971 and 1988. They contain a common region of altered sequence that results in the substitution of six contiguous amino acids (residues 426–431), and by the replacement of Thr 445 by Ala: these changes lead to the decreased affinity of PBP2B for penicillin and hence to increased penicillin resistance (C.D.G. *et al.*, manuscript in preparation). Elsewhere, the genes consist of a mosaic of regions similar to those in sensitive strains ('sensitive

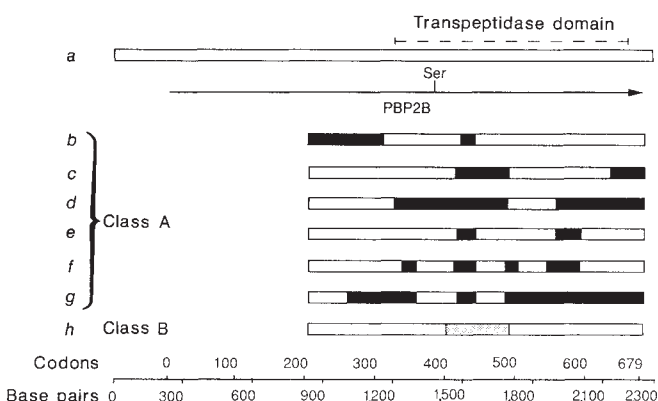


FIG. 1 Mosaic structure of the PBP2B gene in penicillin-resistant strains of *Streptococcus pneumoniae*. *a*, The PBP2B gene of a penicillin-sensitive strain of *S. pneumoniae*. The line terminating in an arrow indicates the coding region. The locations of the penicillin-sensitive transpeptidase domain, and of the active-site serine, are shown. *b–h*, Mosaic structure in the gene from different penicillin-resistant strains (the 5' end of the gene of resistant strains has not been sequenced). The resistant pneumococci represented in *b–g* possess class A PBP2B genes which contain regions (■) that differ at about 14% of sites from those in penicillin-sensitive pneumococci (where these regions overlap in different strains, they differ from one another at <4% of sites). The resistant strain represented in *h* possesses the class B PBP2B gene and contains a region (▨) differing from that of sensitive strains at 21% of sites, but not resembling the diverged regions in the class A PBP2B genes. The unshaded regions are similar in sequence to those in penicillin-sensitive pneumococci (<5% divergence).

blocks') and regions that differ from those in sensitive strains at approximately 14% of sites ('resistant blocks') (Fig. 1b-g). Most of the alterations are synonymous. Where resistant blocks from different class A PBP2B genes overlap, they are similar but not identical (<4% sequence divergence). There is little similarity between the altered regions of class A and class B genes.

We suggest that the identical class B genes derive from a single recombinational event that replaced part of the PBP2B gene with the corresponding region from a related *Streptococcus* that differs by about 21% in sequence from *S. pneumoniae*, and has a homologous PBP with a low affinity for penicillin. The class A genes have also arisen by horizontal gene transfer, but from a different donor species that differs by about 14% in sequence from *S. pneumoniae*. We do not know whether there was a single introduction of DNA from this latter donor species, in which case the different patterns of resistant and sensitive blocks in the class A genes must have arisen by subsequent recombination events, or whether there were several independent introductions of DNA from the donor species.

As a final twist to the story, some strains of the commensal species *S. sanguis* have become resistant to penicillin by acquiring resistant PBP genes from penicillin-resistant *S. pneumoniae*¹³.

A similar picture has emerged from sequence studies of the PBP2 genes of penicillin-resistant strains of *Neisseria meningitidis* and *N. gonorrhoeae*¹⁴⁻¹⁶. In this case, one of the species donating the DNA that confers resistance has been identified as the naturally resistant commensal *N. flavescens*. At least six *Neisseria* 'species', differing in sequence by up to 23%, have exchanged blocks of DNA. Some of the data are shown in Fig. 2.

One case in which there are data both on clonal population structure, provided by electrophoresis, and indicating horizontal gene transfer, concerns another naturally transformable species, *Haemophilus influenzae*¹⁷. Almost all invasive disease is caused by serotype b strains, and the type-specific capsular polysaccharide of these strains is a major factor in virulence. The genes for synthesis contain a central region specific to each of the capsular types, flanked by regions common to all of them. Capsulated strains of *H. influenzae* have a clonal population structure, and related electrophoretic types usually have the same capsular type¹⁸. Strains with capsular type b are, however, found in two distantly related lineages. Restriction maps show many changes between the lineages in the flanking regions, but none for each of 13 restriction enzymes in the serotype-specific central region. Sequencing of parts of the flanking regions from an isolate from each lineage shows 0/250 differences in the central region and 95/795 (12%) differences in the flanking regions. These results imply that serotype switching has occurred by the horizontal transfer of the central, serotype-specific region from a type b strain to a related strain that differs in sequence

by about 12%. This localized recombination event would not alter the electrophoretic type of the recipient.

Transformation and generation of diversity

These examples show that evolution has occurred by the horizontal transfer of chromosomal genes, or parts of genes, and their incorporation into the recipient's chromosome by homologous recombination. There are also cases in which frequent transformation at a locus has a specific function in generating variation. The pili of *N. gonorrhoeae* mediate the binding of gonococci to host epithelia. The major component is the pilin protein, encoded by a single chromosomal gene, *pilE*, or, in some strains, by two closely linked genes. There are also several silent loci, *pilS*, consisting of truncated pilin genes. Changes in antigenic type arise by non-reciprocal recombination between the *pilE* gene and one of the *pilS* genes. This adaptive variation occurs mainly by transformation: that is, by recombination between a *pilS* gene from a dead cell and the *pilE* gene from a living one^{19,20}. Additional variation in gonococcal pilin genes may, presumably, occasionally arise by transformation with DNA from closely related *Neisseria* species that possess *pilS* homologues.

A second example concerns the protection of natural populations of *S. pneumoniae* from bacteriophage²¹. Pneumococci produce either the restriction endonuclease *DpnI* or *DpnII*. These enzymes recognize the same tetranucleotide sequence, but *DpnI* cuts only if it is methylated, *DpnII* only if unmethylated. Phage produced in a *DpnI* strain are thus restricted in a *DpnII* strain, and vice versa. *DpnI* and *DpnII* are encoded by non-homologous genes located at the same position on the chromosome. Pneumococci can switch their restriction enzyme by transformation, involving recombination between homologous regions flanking the *Dpn* genes.

Significance of transformation

These examples of horizontal gene transfer come from three genera which have a special capacity for the uptake and chromosomal incorporation of DNA. Such naturally transformable bacteria are found among both Gram-positive (for example, *Streptococcus*, *Bacillus*) and Gram-negative (*Haemophilus*, *Neisseria*) genera²².

Transformation is probably not an accidental process²³, but has evolved by natural selection. There are two selective advantages that might be obtained by taking in homologous DNA. The new DNA may be used to repair damage to the host chromosome^{24,25}. In addition, a cell may acquire a new gene that confers a selective advantage, as in the evolution of penicillin resistance in *Streptococcus* and *Neisseria*, and in antigenic variation in *N. gonorrhoeae*. Given competence for transformation, we can expect to find evidence of horizontal gene transfer in the form of mosaic genes. Transformation of *N. meningitidis*

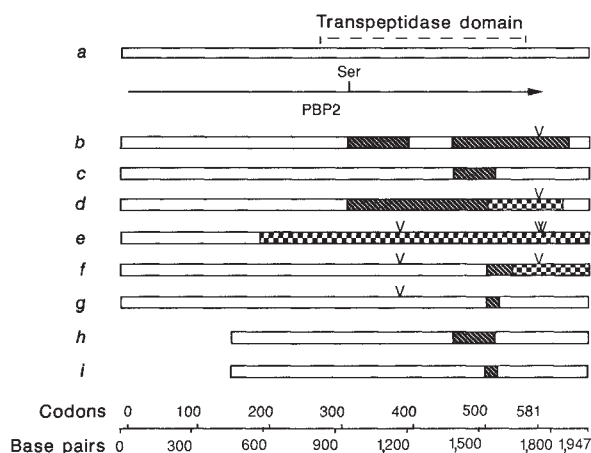


FIG. 2 Mosaic structure of the PBP2 gene in penicillin-resistant *Neisseria* strains. a, The PBP2 gene of a penicillin-sensitive strain of a *Neisseria* species. b-i, PBP2 genes of: penicillin-resistant *N. meningitidis* strains (b-e); penicillin-resistant *N. gonorrhoeae* strains (f,g); a penicillin-resistant *N. lactamica* (h); a penicillin-resistant *N. Polysaccharae* (i). The open regions in b-i represent parts of the PBP2 gene that are almost identical to the corresponding regions in penicillin-sensitive strains of the same species. ■, regions that differ from those in penicillin-sensitive strains at about 23% of sites and have been introduced from *N. flavescens*; ▒, regions that differ from those in penicillin-sensitive strains at about 12% of sites and have been introduced from another *Neisseria* species; V, insertions of additional codons. The 5' end of the PBP2 gene of *N. lactamica* and *N. polysaccharae* has not been sequenced.

and *H. influenzae* has, however, not prevented single clones, identified electrophoretically, from becoming widespread. This is not surprising, because the effects of transformation are local, and would only very occasionally alter electrophoretic type. Local recombination could prevent the development of a clonal population structure, but only if the frequency of such recombination affecting a single gene is high compared with the rate of mutation of that gene.

Horizontal transfer in non-transformable bacteria

What of bacteria that do not show natural transformation, such as *E. coli*? Stoltzfus *et al.*²⁶ present evidence of sequence mosaicism and recombination in a 3,500-base-pair (bp) region close to the *trp* operon. Three strains, ECOR49, 51 and 71, differ by 2–4% in nucleotide sequence, but in one 129-bp region ECOR49 differs by 29% from the other two. DuBose *et al.*²⁷ sequenced the *phoA* gene, coding for alkaline phosphatase, from eight naturally occurring *E. coli* strains. They found 87/1,871 polymorphic sites. These sequences (together with that of the previously sequenced strain, K12) are best explained by a history of recombination, perhaps involving four crossovers. There is also evidence for a mosaic structure in two other gene loci, *trpCBA*²⁸ and *gnd* (ref. 29, and D. E. Dykhuizen and L. Green, unpublished results; for methods of detecting mosaic structure, see refs 30 and 31).

The loci involved in these cases—unlike those in *Neisseria*, *Haemophilus* and *Streptococcus*—have probably not been under strong directional selection. The introduction of a new block of DNA is probably a rare event, and it is therefore not surprising that the mosaic structure is less obvious, because the recombination events have been partially obscured by subsequent point mutation. There are, however, several cases in which horizontal transfer of chromosomal genes has been invoked to explain strongly selected changes in bacteria that are not naturally transformable (for example, antigenic changes in *Salmonella*^{32,33}; enhanced virulence in *E. coli*^{34,35}). There is therefore no reason to believe that horizontal transfer between closely related bacteria is confined in nature to genera that are naturally transformable.

Are there bacterial species?

Bacteria differing in DNA sequence by up to 20% can and do exchange chromosomal DNA. The exchange is usually local, often involving only a few hundred base pairs, and so does not destroy the clonal population structure detected by protein electrophoresis. The evolving population (correspond-

ing to the sexual species in eukaryotes), between whose members genetic exchange is possible, is wider than the named 'species' of the bacteriologist, such as *S. pneumoniae* or *N. meningitidis*.

In eukaryotes, a species has two properties: its members are able to exchange genes with one another, but not with members of other species. Is there anything corresponding to reproductive-isolating mechanisms in prokaryotes? Perhaps the best evidence for genetic isolation between related groups comes from gene sequences in *E. coli* and *Salmonella typhimurium*. Sharp and Li³⁶ found that the divergence between different loci can be explained by different constraints on codon usage: loci with highly biased codon usage have diverged less. This is to be expected if the two genera have diverged through the accumulation of selectively neutral alleles. But, the correlation between codon usage and divergence would be destroyed if there was frequent gene exchange. Therefore the data imply that there has been little or no transfer of chromosomal genes between the two taxa.

If this pattern of isolation between related groups is typical, then units analogous to sexual species, in respect both of gene exchange within, and of isolation between, populations, may indeed exist in prokaryotes. Several mechanisms could limit the exchange of chromosomal DNA between related groups of bacteria. Restriction enzymes are unlikely to be a serious barrier to localized recombination. In *S. pneumoniae*, they seem to have evolved so as to protect against phage infection without affecting conjugal DNA transfer or transformation³⁷. Even where restriction enzymes operate, as in phage-mediated transduction, the resulting DNA fragments may still participate in localized recombination. A more serious barrier to genetic exchange may be the repair systems that recognize and abort heteroduplexes formed between DNA strands containing extensive mismatching³⁸.

We conclude that although most bacteria have a clonal population structure, individual genes may be mosaics arising from local gene exchange between related species. Some of the examples concern changes that have been driven by strong selection, in species that show natural transforming ability. But even with little selection and in taxa not competent for transformation, there is evidence of localized sex. In some bacteria at least, evolving populations are wider than named 'species'. □

John Maynard Smith, Christopher G. Dowson and Brian G. Spratt are at the School of Biological Sciences, University of Sussex, Brighton BN1 9QG, UK.

- Griffith, F. J. *J. Hyg. Camb.* **27**, 113–159 (1928).
- Selander, R. K. & Levin, B. R. *Science* **210**, 545–547 (1980).
- Orskov, F. & Orskov, I. *J. infect. Dis.* **148**, 346–357 (1983).
- Whittam, T. S., Ochman, H. & Selander, R. K. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1751–1755 (1983).
- Selander, R. K., Caugant, D. A. & Whittam, T. S. in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology* (ed. Neidhardt, F. C.) 1625–1648. (American Society for Microbiology, Washington, 1987).
- Musser, J. M., Granoff, D. M., Pattison, P. E. & Selander, R. K. *Proc. natn. Acad. Sci. U.S.A.* **82**, 5078–5082 (1985).
- Caugant, D. A. *et al. J. Bacteriol.* **169**, 2781–2792 (1985).
- Olyhoek, T., Crowe, B. A. & Achtman, M. *Rev. Infect. Dis.* **9**, 665–692 (1987).
- Caugant, D. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 4927–4931 (1986).
- Spratt, B. G. & Cromie, K. D. *Rev. Infect. Dis.* **10**, 699–711 (1988).
- Spratt, B. G. in *Handbook of Experimental Pharmacology* Vol. 91 (ed. Bryan, L. E.) 77–100 (Springer-Verlag, Berlin, 1989).
- Dowson, C. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 8842–8846 (1989).
- Dowson, C. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **87**, 5858–5862 (1990).
- Spratt, B. G. *Nature* **332**, 173–176 (1988).
- Spratt, B. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 8988–8992 (1989).
- Lujan Marca, R., Zhang, Q.-Y., Sáez Nieto, J. A., Jones, D. M. & Spratt, B. G. *Antimicrob. Agents Chemother.* (in the press).
- Kroll, J. S. & Moxon, E. R. *J. Bact.* **172**, 1374–1379 (1990).
- Musser, J. M., Kroll, J. S., Moxon, E. R. & Selander, R. K. *Proc. natn. Acad. Sci. U.S.A.* **85**, 7758–7762 (1988).
- Seifert, H. S., Ajioka, R. S., Marchal, C., Sparling, P. F. & So, M. *Nature* **336**, 392–395 (1988).
- Gibbs, C. P. *et al. Nature* **338**, 651–652 (1989).
- Lacks, S. A., Mannarelli, B. M., Springhorn, S. S. & Greenberg, B. *Cell* **46**, 993–1000 (1986).
- Stewart, G. J. & Carlson, C. A. *Rev. Microbiol.* **40**, 211–235 (1986).
- Levin, B. R. in *The Evolution of Sex* (eds Michod, R. E. & Levin, B. R.) 194–211 (Sinauer, Sunderland, Massachusetts 1988).
- Michod, R. E., Wojciechowski, M. F. & Hoelzer, M. A. *Genetics* **118**, 31–39 (1988).
- Wojciechowski, M. F., Hoelzer, M. A. & Michod, R. E. *Genetics* **121**, 411–422 (1989).
- Stoltzfus, A., Leslie, J. F. & Milkman, R. *Genetics* **120**, 345–358 (1988).
- DuBose, R. F., Dykhuizen, D. E. & Hartl, D. L. *Proc. natn. Acad. Sci. U.S.A.* **85**, 7036–7040 (1988).
- Milkman, R. & Crawford, I. P. *Science* **221**, 378–380 (1983).
- Dykhuizen, D. E. & Green, L. *Genetics* **113**, 871 (1986).
- Stephens, J. C. *Molec. Biol. Evol.* **2**, 539–556 (1985).
- Sawyer, S. *Molec. Biol. Evol.* **6**, 526–538 (1989).
- Beltran, P. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 7753–7757 (1988).
- Smith, N. H., Beltran, P. & Selander, R. K. *J. Bact.* **172**, 2209–2216 (1990).
- Plos, K. *et al. Infect. Immun.* **57**, 1604–1611 (1989).
- Arthur, M. *et al. Infect. Immun.* **58**, 471–479 (1990).
- Sharp, P. M. & Li, W.-H. *Molec. Biol. Evol.* **4**, 222 (1987).
- Cerritelli, S., Springhorn, S. S. & Lacks, S. A. *Proc. natn. Acad. Sci. U.S.A.* **86**, 9223–9227 (1989).
- Rassayguier, C., Thaler, D. S. & Radman, M. *Nature* **342**, 396–401 (1989).

ACKNOWLEDGEMENTS. Work from our laboratory was funded by the Wellcome Trust and the Medical Research Council. B.G.S. is a Wellcome Trust Principal Research Fellow.

The density field of the local Universe

Will Saunders^{*†}, Carlos Frenk[‡], Michael Rowan-Robinson[†], George Efsthathiou^{*},
Andy Lawrence[§], Nick Kaiser^{||}, Richard Ellis[‡], John Crawford[†], Xiao-Yang Xia[‡]
& Ian Parry[‡]

^{*} Department of Astrophysics, University of Oxford, Keble Road, Oxford OX2 3RJ, UK

[†] Astronomy Unit, Queen Mary and Westfield College, Mile End Road, London E1 4NS, UK

[‡] Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK

[§] Department of Physics, Queen Mary and Westfield College, Mile End Road, London E1 4NS, UK

^{||} Canadian Institute for Theoretical Astrophysics, University of Toronto, Toronto, Ontario, Canada, M5S 1A1

An all-sky redshift survey of galaxies detected by IRAS (the Infrared Astronomical Satellite) has been used to map the Universe out to $140h^{-1}$ Mpc (the Hubble constant $H_0 \equiv 100h \text{ km s}^{-1} \text{ Mpc}^{-1}$). Well-known superclusters and voids are seen, as are others not previously identified. The inferred underlying distribution of density is found to be skewed to high densities (the voids are larger than the superclusters but depart less from the mean density); and there is more structure on large scales than is predicted by the standard cold dark matter theory of galaxy formation.

THE IRAS survey allows the construction of cleanly selected all-sky galaxy catalogues, of greater solid angle and depth than previously existing optically based catalogues. It is thus ideal for mapping and quantifying the large-scale structure of the Universe. We have recently completed the QDOT (Queen Mary and Westfield-Durham-Oxford-Toronto, *née* QCD) redshift survey, consisting of 2,163 randomly sampled¹ IRAS galaxies brighter than 0.6 Jy, covering virtually all of the sky at galactic latitudes $|b| > 10^\circ$. Redshifts for 56% of the sample were obtained in 1986–88 using Faint Object Spectrographs on the Isaac Newton, William Herschel and the Anglo-Australian telescopes, the identified sources ranging from 15 mag to 20 mag. Published redshifts were used for 34% of the sources, and redshifts for a further 7% were supplied by other workers before publication. Of the 3% without redshifts, 1.2% are thought to be galactic, 1.3% are galaxies and 0.4% (8 sources) are blank fields at the POSS (Palomar Observatory Sky Survey) limit of ~ 20.5 mag.

Our overall completeness, excluding galactic sources, is then 98%. The aim of this survey is to determine the strength of large-scale clustering, and the velocity field and cosmography of the local Universe. The source selection for the two-dimensional parent catalogue and the redshift survey itself will be described in forthcoming papers (ref. 2 and A.L. *et al.*, manuscript in preparation). The 60- μm luminosity function and evolutionary properties of the sample³, the inferred velocity field and the origin of our motion with respect to the microwave background (ref. 4 and N.K. *et al.*, manuscript in preparation), and another approach to measuring large-scale clustering⁵ are discussed elsewhere. The sky distribution of the survey is shown in Fig. 1. A complementary, shallower but fully sampled, all-sky IRAS redshift survey has been carried out by Strauss and co-workers⁶.

Here we describe the distribution of galaxies out to $140h^{-1}$ Mpc. We formulate a new statistical technique for quantifying large-scale structure and present the results of this method and a comparison with the standard cold dark matter theory of galaxy formation.

Previous surveys

Rigorous work on cosmography and the statistics of large-scale structure in the distribution of galaxies have been hampered by the lack of cleanly selected survey material. Although superclustering was observed by Herschel, our knowledge of superclustering is still largely qualitative. Large structures were known from the two-dimensional catalogues of Shapley and Ames⁷ and Zwicky⁸, and the Shane and Wirtanen⁹ counts, but the redshift surveys pioneered by Davis *et al.*¹⁰ have transformed our understanding of galaxy clustering. Several large redshift surveys are being done or have recently been completed^{6,11–14}, but none has the combination of sky coverage and depth available here. Abell clusters have been used to search for structure on a very much larger scale^{15–17}.

FIG. 1 Sky distribution of galaxies in the QDOT survey. Also shown are the areas not included in the survey owing to incomplete satellite coverage, source confusion or redshift incompleteness. Symbol size is scaled as $\log(1/V)$.

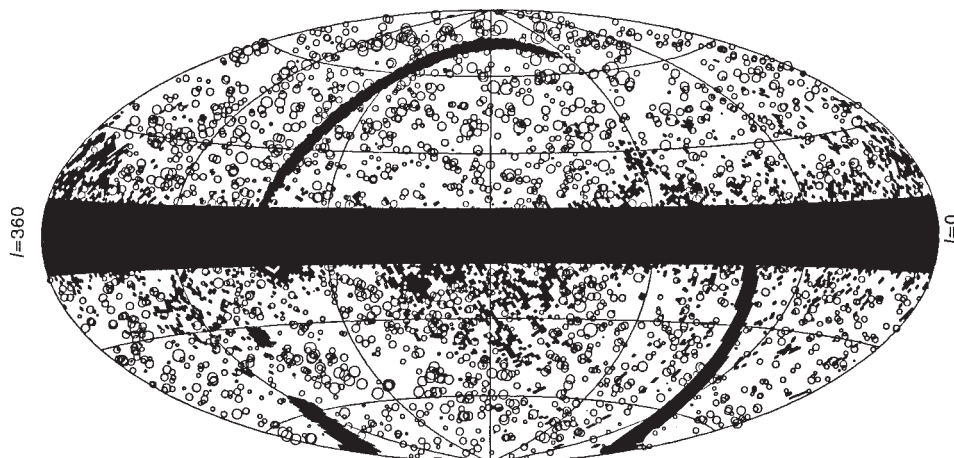


TABLE 1 Density maxima and minima from the QDOT survey

				X	Y	Z	Confidence			
Name		<i>l</i>	<i>b</i>	Velocity (km s ⁻¹)	(h ⁻¹ Mpc)			Density	Noise	level (%)
Superclusters, $r_{\text{sm}}=5h^{-1}$ Mpc										
Vi	Virgo	269	67	1,089	0	-4	10	4.04	0.20	<0.01
EF	Eridanus-Fornax	228	-54	1,148	-5	-5	-9	2.48	0.21	<0.01
Hy	Hydra	268	29	3,401	-1	-30	16	4.36	0.40	<0.01
Ce	Centaurus	304	13	2,764	15	-22	6	2.98	0.41	<0.01
NC	North Centaurus	314	31	4,156	25	-25	21	3.03	0.44	0.01
S1	New supercluster	218	-15	2,525	-19	-15	-6	2.63	0.41	0.06
Voids, $r_{\text{sm}}=5h^{-1}$ Mpc										
SV1	SSRS void 1	259	-58	3,199	-3	-17	-27	0.00	0.35	0.03
LV	Local void	48	-17	2,326	15	16	-7	0.00	0.36	0.04
V3	New void	69	-81	3,336	2	5	-33	0.00	0.38	0.10
V1	New void	143	-17	1,751	-13	10	-5	0.23	0.31	0.16
V2	New void	210	13	2,829	-24	-14	6	0.01	0.42	0.35
SV2	SSRS void 2	26	-38	4,027	29	14	-25	0.02	0.43	0.41
Superclusters, $r_{\text{sm}}=10h^{-1}$ Mpc										
VHC	Virgo-Hydra-Centaurus	289	26	2,429	7	-21	11	1.86	0.13	<0.01
PP	Perseus-Pisces	149	-9	5,949	-50	30	-9	3.08	0.36	<0.01
N1	N1600	194	-24	4,406	-39	-10	-18	2.17	0.22	<0.01
PI	Pavo-Indus	347	-20	5,228	48	-11	-18	2.26	0.24	<0.01
S2*	(140-33)	142	-38	9,906	-62	48	-61	4.08	0.47	<0.01
S3	New supercluster	273	-44	8,354	3	-60	-58	2.88	0.37	<0.01
A2*	A2197/2199	60	38	~10,000	40	70	64	5.49	0.49	<0.01
HX†	Hercules extention	13	53	~9,000	51	11	70	4.14	0.38	<0.01
Co†	Coma	71	71	~7,400	8	23	70	2.18	0.31	0.02
A1†	A1367	220	79	~7,100	-10	8	70	2.27	0.32	0.02
S6†	New supercluster	222	-33	~11,000	-70	-63	-62	3.33	0.57	0.03
S4	New supercluster	281	0	6,120	12	-60	0	2.02	0.33	0.37
S5	New supercluster	96	-33	7,386	-6	62	-40	1.97	0.32	0.46
Voids, $r_{\text{sm}}=10h^{-1}$ Mpc										
SV	SSRS voids 1, 2	15	-78	3,568	7	2	-35	0.41	0.15	<0.01
LV	Local void	20	10	4,643	43	15	8	0.15	0.26	<0.01
V4	New void	62	35	4,904	19	36	28	0.31	0.20	<0.01
EV	Eridanus void	303	-39	6,289	27	-41	-40	0.26	0.26	0.04
CV1	Cfa 14.5+60	113	54	7,410	-17	40	60	0.28	0.31	0.18
CV	Coma void	36	54	6,151	29	21	50	0.36	0.24	0.23
T3	Tully void 3	239	8	7,064	-36	-60	10	0.09	0.43	0.49
T2‡	Tully void 2	47	-45	6,940	33	36	-49	0.28	0.30	0.58
Superclusters, $r_{\text{sm}}=20h^{-1}$ Mpc										
He	Hercules	16	56	11,319	60	18	94	2.47	0.21	<0.01
A2§	A2197/99	55	40	9,769	43	62	62	1.98	0.17	<0.01
S8†	New supercluster	162	0	~15,000	-140	46	0	2.94	0.48	0.05
S2§	(140-33)	137	-41	11,612	-64	60	-76	1.81	0.22	0.07
NH	Near Horologium	255	-31	15,557	-35	-129	-80	2.37	0.37	0.07
S7	New supercluster	11	-35	13,821	110	22	-80	2.07	0.29	0.09
S9†	New supercluster	61	-46	~19,500	65	118	-140	3.16	0.64	0.38
Voids, $r_{\text{sm}}=20h^{-1}$ Mpc										
BV	Bootes void	81	52	15,713	16	94	125	0.04	0.37	0.06

(*l*, *b*) are the galactic coordinates. Density is the peak density divided by the mean. Noise is the r.m.s. Poisson shot noise for a uniform underlying density field. Confidence level is the probability of the feature being noise. Total mass associated with features varies as Density $\times r_{sm}^3$.

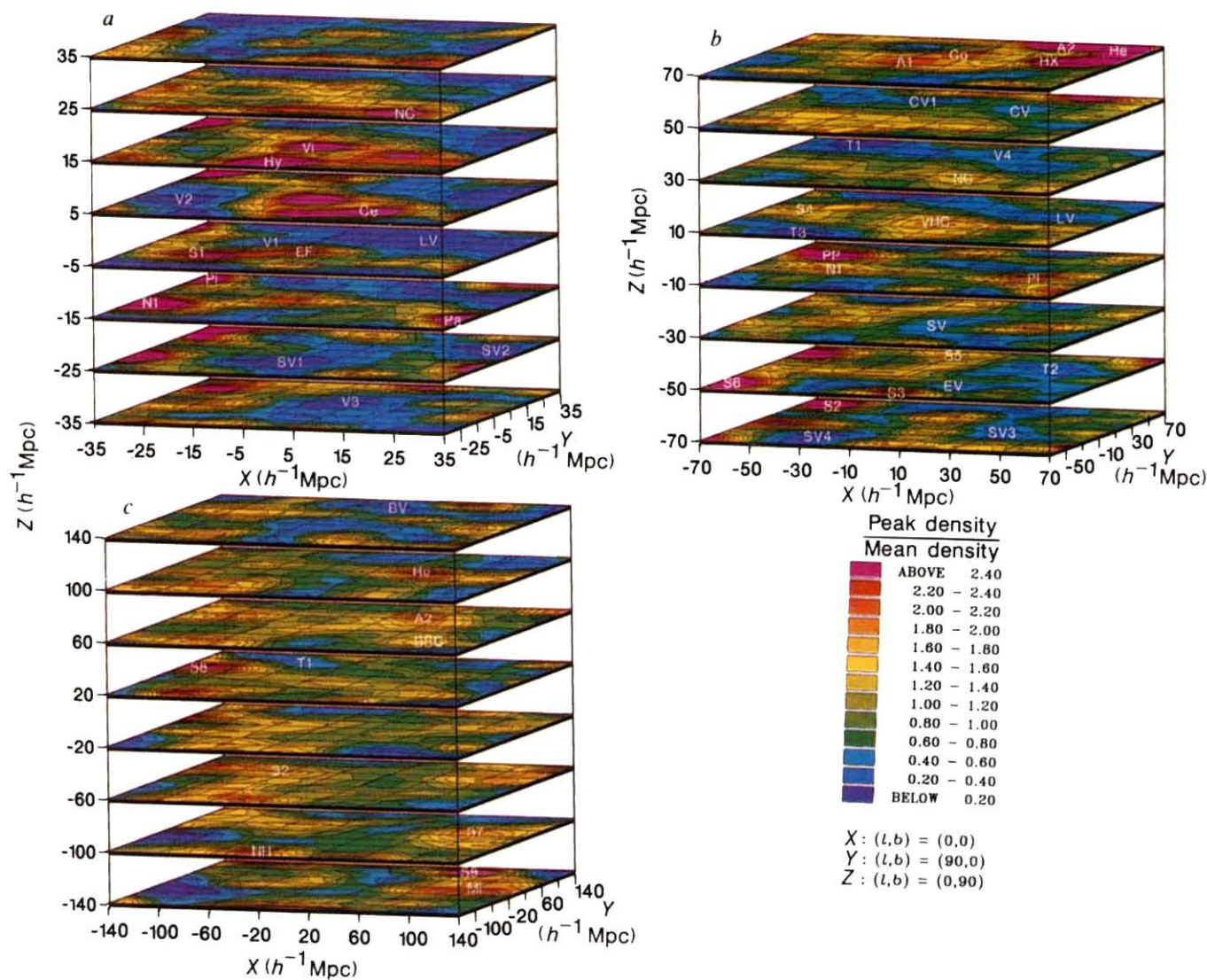
* Also seen at $20h^{-1}$ Mpc resolution. † Just off edge of the map. ‡ Low signal-to-noise but an extended feature. § Also seen at $10h^{-1}$ Mpc resolution.

Clustering on small scales ($\leq 10h^{-1}$ Mpc) is reasonably well understood, mostly by the methods developed by Peebles and co-workers¹⁸. On larger scales, systematic effects in the measurement of small amplitude fluctuations have made progress very difficult. The IRAS survey combines large sky coverage, great depth and negligible extinction: we can hope to see very large but low-amplitude features reliably.

Constructing the density field

Our aim is to find the smoothed density field of the galaxies in our redshift survey, both to identify major features and to study the statistics of the density field. We want to use all of the available information to estimate the density field. In any flux-limited redshift survey, the observed number density of galaxies falls off with distance (because they have to be increasingly luminous to be seen above our flux limit). To produce a volume-

limited sample would be very inefficient, especially for IRAS galaxies, which have a broad luminosity function (IRAS galaxies are seen with a very wide range of luminosities). This means we must weight the observed distribution by the inverse of the selection function, defined to be the expected number density of galaxies at given distance luminous enough to be included in our sample, and we must be sure we can do this reliably. The luminosity function for IRAS galaxies is accurately known³ and so this is straightforward. IRAS galaxies also show strong evolution^{3,19,20} in the sense that there are more galaxies of given luminosity seen at large distances (and hence look-back time) than expected from their number density nearby, and this has a significant effect on the selection function. We used a selection function calculated from solution 19 of ref. 3, which includes density evolution at a rate $(1+z)^{6.7}$. We assumed an Einstein-deSitter universe throughout.



a
NC North Centaurus

Vi Virgo
Hy Hydra

Ce Centaurus
V2 new void

LV Local void
EF Eridanus–Fornax
S1 new supercluster
V1 new void

N1 N1600
Pi Pisces
Pa Pavo

SV1 SSRS void 1
SV2 SSRS new void

V3 new void

b
Co Coma
A1 A1367
He Hercules
A2 A2197/2199
HX Hercules extension

CV1 Cfa 14.5+60
CV Coma void

NC North Centaurus
V4 new void
T1 Tully void 1

VHC Virgo–Hydra–Centaurus
LV Local void
S4 new supercluster
T3 Tully void 3

PP Perseus–Pisces
N1 N1600
PI Pavo–Indus

SV SSRS voids 1,2

EV Eridanus void
T2 Tully void 2
S3 new supercluster
S4 new supercluster
S5 new supercluster

SV4 SSRS void 4
SV3 SSRS void 3
S2 (140–33)

c
BV Bootes void

He Hercules

A2 A2197/2199

SSC Shapley's supercluster

B9 new supercluster
T1 Tully void 1

S2 (140–33)

NH Near Horologium
S7 new supercluster

Mi Microscopium
S9 new supercluster

TABLE 2 Moments of galaxy density

	5h ⁻¹ Mpc	Smoothing scale 10h ⁻¹ Mpc	20h ⁻¹ Mpc
QDOT			
$\langle(d\rho/\rho)^2\rangle$	0.436 ± 0.091	0.184 ± 0.050	0.0669 ± 0.019
$\langle(d\rho/\rho)^3\rangle$	0.393 ± 0.209	0.084 ± 0.061	0.0250 ± 0.017
CDM			
$\langle(d\rho/\rho)^2\rangle$	0.447 ± 0.018	0.131 ± 0.005	0.0192 ± 0.0013
$\langle(d\rho/\rho)^3\rangle$	0.358 ± 0.037	0.035 ± 0.005	-0.0010 ± 0.0004

CDM results are the averages of 30 simulations. The uncertainties are quoted at the 1 σ level.

To find a continuous density field, we have to smooth the observed point distribution of galaxies, and we have used a gaussian smoothing function to do this. This smoothing should be borne in mind when examining the density maps—structure on scales below the smoothing scale is lost. The smoothing scale required at great distances (where relatively few galaxies are luminous enough to be included in our sample) is much larger than is necessary locally, so we present our results using a series of smoothing scales, and in each case the map extends as far as discreteness noise allows. Each map is then an ‘expanded view’ of the central eighth of the next one.

The survey covers 74% of the sky. Regions not included are the coverage gaps in the original IRAS survey, or where confusion due to galactic sources makes source extraction unreliable, or at galactic latitude $|b| < 10^\circ$ where identification and redshift acquisition became difficult. The resulting mask is shown in Fig. 1. To take this mask into account when calculating the density field, we have calculated how much of any given gaussian sphere has been excluded by the mask.

The distances for the galaxies are taken directly from their redshifts, except for a small number of nearby galaxies (all with velocity < 500 km s⁻¹) and those in the core of Virgo. For these galaxies, peculiar velocities are comparable to their Hubble velocities, and we have assigned them pseudo-redshifts derived from direct distance indicators^{3,21}. Apart from this, we have

made no correction to the redshifts for peculiar motions. It has been shown²² that this distorts the density field and leads to slight overestimation of the density contrast; we will investigate the velocity field and its effect on observed clustering elsewhere. Here it is sufficient that the effect applies equally to the QDOT survey and the *N*-body simulations with which we will compare it.

We calculate the smoothed density d_i on a lattice of points $\mathbf{r}_i$ as follows. Using the luminosity function we calculate the selection function $\psi(\mathbf{r})$, that is, the expected number density of galaxies bright enough to be seen above our flux limit of 0.6 Jy at distance $|\mathbf{r}|$ for a uniform density field without clustering.

We define $w_i(\mathbf{r})$ to be the gaussian smoothing function centred on $\mathbf{r}_i$ with radius r_{sm} :

$$w_i(\mathbf{r}) = \frac{1}{(2\pi r_{\text{sm}}^2)^{3/2}} \exp\left(-\frac{1}{2} \frac{|\mathbf{r}_i - \mathbf{r}|^2}{r_{\text{sm}}^2}\right) \quad (1)$$

Then we have

$$d_i = \frac{\sum_j w_i(\mathbf{r}_j) / \psi(\mathbf{r}_j)}{\int_V w_i(\mathbf{r}) dV} \quad (2)$$

where the sum is over the observed distribution of galaxies at positions $\mathbf{r}_j$, the integral is over the unmasked volume, and both are truncated at $3r_{\text{sm}}$ from $\mathbf{r}_i$. We calculate the smoothed density field on a 15³ grid, with a grid interval equal to the smoothing radius r_{sm} —this means the volume is heavily oversampled.

Map of the local Universe

Figure 2a–c shows the density field in cubes of different sizes, with smoothing scales 5, 10 and 20h⁻¹ Mpc, respectively. In each case our galaxy is at the centre, the *X* axis is towards the Galactic Centre (l, b) = (0, 0), the *Y* axis is towards (l, b) = (90, 0), and the *Z* axis is towards the north galactic pole. The slices displayed are alternate 15² planes of the whole 15³ grid, so the separation between slices is $2r_{\text{sm}}$ and they are nearly independent.

The noise in the maps is a function of position. At the centre of each map, the noise is < 1 contour level. At the edges and corners of the maps, the noise is typically 2.5 and occasionally 4 contour levels, so all but the most extreme or extensive features seen here should be treated with caution. In Table 1, we list the highest signal-to-noise maxima and minima in the maps; each is significant at the 99% level, assuming Poisson statistics. Although there are 15³ grid points, the gaussian spheres overlap and hence oversample the density field 43-fold (see below). There are then effectively ~ 80 independent volumes in each grid, and we expect ≤ 1 spurious feature per map at this confidence level. There are also features at or slightly beyond the edges of the maps; these are not included in Table 1, but are labelled in the maps where they correspond to well-known structures.

One of the most striking features of the maps is their connectedness. Both underdense and overdense regions link up with each other, and there is no obvious difference in the topology of overdense and underdense regions. A more detailed topological investigation will be discussed elsewhere (Moore, B. E. *et al.*, manuscript in preparation).

Moments of the density distribution

The grids of densities used to generate the maps contain information about the statistical properties of the large-scale distribution of matter in the Universe. In Fig. 3 we show the histograms of observed densities for the three smoothing scales. These histograms have been broadened and skewed by the effects of Poisson shot noise—also shown are histograms for Poisson simulations of the data (with the same number of galaxies, selection function and mask), that is, pure shot noise. The difference between the Poisson and data histograms is due to clustering, and we would like to quantify this difference. We derive expressions for the

◀ FIG. 2 Density maps for cubical volumes of half-side 35, 70 and 140h⁻¹ Mpc. Our galaxy is in the centre of each cube, and the *X*, *Y*, *Z* axes are in the directions (l, b) = (0, 0), (l, b) = (90, 0), (l, b) = (0, 90), respectively. The noise varies from less than one contour level at the centre to 2.5–4 contour levels at the edges and corners. a, Half-side is 35h⁻¹ Mpc, $r_{\text{sm}} = 5h^{-1}$ Mpc. Map is dominated by the various components of the supergalactic plane—Virgo, Hydra, Centaurus, North Centaurus and Eridanus–Fornax. There is a bridge of galaxies (S1) from Eridanus–Fornax to N1600. The nearer part of Pisces is seen, and extends through the galactic plane. Voids 1 and 2 (SV1, SV2) noted in the SSRS (southern sky redshift survey)¹¹ are clearly seen. There are new voids V1 (an extension of the Local void) and V2 at low galactic latitudes, and V3, an extension of SV1. The Local void is very extensive, and links up with SV2 and V1. b, Half-side is 70h⁻¹ Mpc, $r_{\text{sm}} = 10h^{-1}$ Mpc. Shows the whole supergalactic plane. At this resolution, the Local supercluster (Virgo, Hydra and Centaurus) forms a single overdense structure, resolved from Pavo and Perseus–Pisces. There is not much evidence for the ‘Great Attractor’, which would be at about [25, –33, 13] (ref. 29). The Coma/A1367 supercluster is at the top and forms part of the ‘Great Wall’¹². There is a link from here to North Centaurus 80h⁻¹ Mpc away. The nearer portions of the Hercules supercluster show up dramatically, and there is a clear extension of the supercluster towards higher galactic latitude. There are new superclusters S2 (140 – 33 of ref. 4), S3, S4, S5 and S6. Also seen are SSRS voids 1/2, 3 and 4 (ref. 11), Tully’s voids 1, 2 and 3 (ref. 16), and the Eridanus void³⁰. c, Half-side 140h⁻¹ Mpc, $r_{\text{sm}} = 20h^{-1}$ Mpc. A2197/9 and Hercules superclusters are seen again, extending to 12,000 km s⁻¹. Nearer Horologium is seen, and part of Further Horologium³⁰. The Microscopium/Further Indus supercluster³⁰ is also seen, which together with S9 forms Tully’s Aquarius–Capricornus supercluster complex. The band of overdensity running down the left-hand side and including S8 is Tully’s Pisces–Cetus complex¹⁶. ‘Shapley’s supercluster’³¹ is the (very weak) feature SSC. The only high signal-to-noise void in the cube is Bootes, but there are clearly other very large voids at the edges of the cube.

moments of the underlying density distribution (that which when sampled in the same way as the real data gives rise to the observed density distribution). An alternative selection-function-independent approach is presented elsewhere⁵.

We assume that galaxies arise as a Poisson process from some underlying galaxy density field $\rho_g(\mathbf{r})$ —this is the Poisson model described by Peebles¹⁸. According to current ideas, ρ_g may be a biased measure of the mass density field ρ_m , with $\delta\rho_g/\rho_g = b(\delta\rho_m/\rho_m)$. Recent determinations of the bias parameter appropriate to IRAS galaxies (assuming an Einstein-deSitter universe) have given b_{IRAS} in the range 1.2–1.6 (refs 4, 23, 24 and N.K. *et al.*, manuscript in preparation). We also assume that the shape (but not the normalization) of the luminosity function is the same everywhere. Then the probability of there being a galaxy with luminosity in the range $[L, L + \delta L]$ in a volume element δV at $\mathbf{r}$ is $\rho_g(\mathbf{r})\phi(L)\delta L\delta V$, where $\phi(L)$ is the luminosity function. We now drop the subscript g and density will always refer to the galaxy density ρ_g unless otherwise specified.

Let $\hat{\rho}_i$ be the underlying (galaxy) density field at a point $\mathbf{r}_i$ once it has been smoothed with the gaussian smoothing function w_i from equation (1)

$$\hat{\rho}_i = \frac{\int_V w_i(\mathbf{r})\rho(\mathbf{r})dV}{\int_V w_i(\mathbf{r})dV} \quad (3)$$

We have a grid of observed gaussian smoothed densities d_i . Each d_i is an unbiased estimate of $\hat{\rho}_i$, and we will use the observations d_i to estimate the moments of $\hat{\rho}$.

There are three parts to our strategy. (1) We need to quantify the properties of the shot noise for each grid point. This depends on $\hat{\rho}_i$, which is unknown. We define the random variable $\tilde{d}_i$ to be possible outcomes for the observed densities once $\hat{\rho}_i$ has been specified, and find the moments of $\tilde{d}_i$ as functions of $\hat{\rho}_i$. (2) The moments of the possible outcomes of d_i are uniquely

specified in terms of the moments of $\hat{\rho}$ and $\tilde{d}_i$, and we find these relationships. (3) We use these relationships to form minimum-variance estimators for the moments of $\hat{\rho}$. For brevity, we drop the subscript i and argument $\mathbf{r}$.

First we need to find, for each grid point, the moments of $\tilde{d}$. Let $\tilde{\delta}$ be the number of galaxies in an infinitesimal volume element δV at $\mathbf{r}$. Then the probability of finding a galaxy in δV is $\delta P = \rho\psi\delta V$, and the moment generating function for $\tilde{\delta}$ is

$$M_{\tilde{\delta}}(t) = 1 + \rho\psi\delta V(e^t - 1) \quad (4)$$

where t is a dummy variable. The observed density $\tilde{d}$ will be the selection and smoothing-function-weighted sum of the $\tilde{\delta}$ s

$$\tilde{d} = \sum \frac{w}{\psi} \tilde{\delta} \quad (5)$$

and will have the moment generating function

$$M_{\tilde{d}}(t) = \Pi M_{\tilde{\delta}}\left(\frac{w}{\psi}t\right) = \exp\left[\hat{\rho} \int_V \psi(e^{wt/\psi} - 1)dV\right] \quad (6)$$

We have made a further assumption here: we have assumed that the moments of $\tilde{d}$ are the same as they would be if we had a constant density $\hat{\rho}$ across the sphere. This is true for a purely Poisson process, that is, for a uniform selection function and top-hat weighting²⁵, and should not introduce any serious error. Note that once a density field $\rho(\mathbf{r})$ has been laid down, the $\tilde{\delta}$ s are independent.

It is easier to work with central moments, with corresponding moment generating function

$$M_{\tilde{d}-\hat{\rho}}(t) = \exp\left[\hat{\rho}\left(\int_V \psi(e^{wt/\psi} - 1)dV - t\right)\right] \quad (7)$$

It is also convenient to introduce the quantity $\Sigma_m = \int_V \psi(w/\psi)^m dV$. Differentiation of $M_{\tilde{d}-\hat{\rho}}(t)$ then gives the central moments of $\tilde{d}$ to be

$$\begin{aligned} E(\tilde{d} - \hat{\rho}) &= 0 & E(\tilde{d} - \hat{\rho})^2 &= \hat{\rho}\Sigma_2 \\ E(\tilde{d} - \hat{\rho})^3 &= \hat{\rho}\Sigma_3 & E(\tilde{d} - \hat{\rho})^4 &= 3(\hat{\rho}\Sigma_2)^2 + \hat{\rho}\Sigma_4 \\ E(\tilde{d} - \hat{\rho})^5 &= 10\hat{\rho}^2\Sigma_2\Sigma_3 + \hat{\rho}\Sigma_5 \\ E(\tilde{d} - \hat{\rho})^6 &= 15(\hat{\rho}\Sigma_2)^3 + 15\hat{\rho}^2\Sigma_2\Sigma_4 + 10(\hat{\rho}\Sigma_3)^2 + \hat{\rho}\Sigma_6 \end{aligned} \quad (8)$$

$\Sigma_2, \dots, \Sigma_6$ are calculated for each grid point by direct numerical integration over the unmasked portion of the gaussian sphere.

Next, we need the relationships between the moments of $\tilde{d}$, $\hat{\rho}$ and d above. Expanding $(d - 1)^n$ about $\hat{\rho}$ gives

$$\begin{aligned} E(d - 1)^n &= E((d - \hat{\rho}) + (\hat{\rho} - 1))^n \\ &= E_{d,\hat{\rho}}\left[\sum_{m=1}^n {}^nC_m(\tilde{d} - \hat{\rho})^m(\hat{\rho} - 1)^{n-m}\right] \end{aligned} \quad (9)$$

where $E_{d,\hat{\rho}}$ means the expectation for all d and all $\hat{\rho}$, taken first over $\tilde{d}$ for fixed $\hat{\rho}$ and then over $\hat{\rho}$.

Using the equations (8) above gives

$$E(d - 1)^n = R_n + X_n \quad (10)$$

where $R_n = E(\hat{\rho} - 1)^n$ is the n th central moment of $\hat{\rho}$, and X_n is the term due to shot noise in the n th moment of d

$$\begin{aligned} X_1 &= 0 & X_2 &= \Sigma_2 & X_3 &= \Sigma_3 + 3R_2\Sigma_2 \\ X_4 &= \Sigma_4 + 3(R_2 + 1)\Sigma_2^2 + 4R_2\Sigma_3 + 6(R_2 + R_3)\Sigma_2 \\ X_5 &= \Sigma_5 + 10(R_2 + 1)\Sigma_2\Sigma_3 + 15(2R_2 + R_3)\Sigma_2^2 \\ &\quad + 5R_2\Sigma_4 + 10(R_2 + R_3)\Sigma_2 \\ X_6 &= \Sigma_6 + 15(R_3 + 3R_2 + 1)\Sigma_2^3 + 15(R_2 + 1)\Sigma_2^2 \\ &\quad + 60(R_3 + 2R_2)\Sigma_2\Sigma_3 + 6R_2\Sigma_5 \\ &\quad + 45(R_2 + 2R_3 + R_4)\Sigma_2^2 + 15(R_2 + R_3)\Sigma_4 \\ &\quad + 20(R_3 + R_4)\Sigma_3 + 15(R_5 + R_4)\Sigma_2 \end{aligned} \quad (11)$$

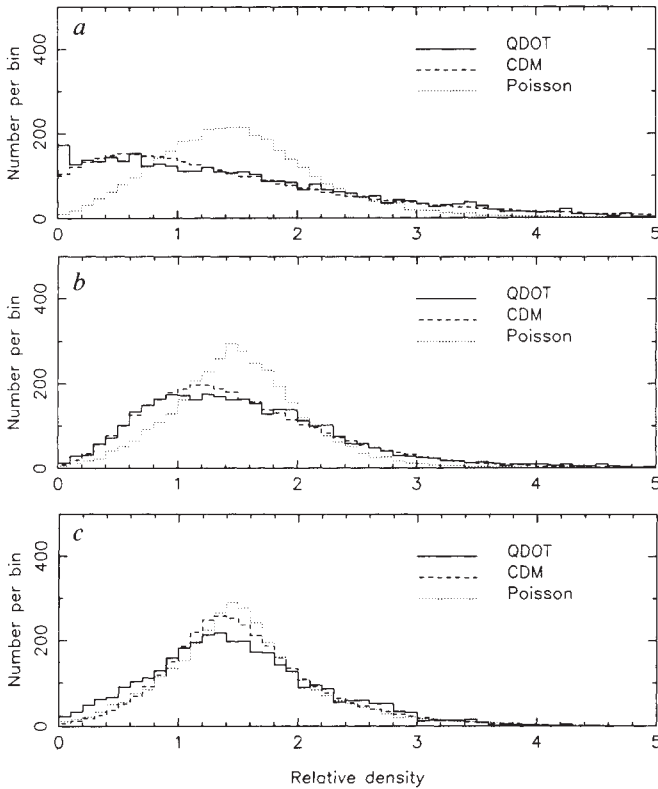


FIG. 3 Histograms of densities for *a*, $r_{\text{sm}} = 5h^{-1}$ Mpc, *b*, $r_{\text{sm}} = 10h^{-1}$ Mpc and *c*, $r_{\text{sm}} = 20h^{-1}$ Mpc, compared with Poisson and CDM simulations of the data. For the simulations, the histograms are averages over many realizations.

We are now in a position to form minimum-variance estimators for the moments of $\hat{\rho}$. We need to reintroduce the subscripts i . For each grid point i and for each n , $(d_i - 1)^n - X_{i,n}$ is an unbiased estimator of R_n , the n th central moment of $\hat{\rho}$. Our final estimator for R_n will then be a weighted mean of the estimates from each grid point

$$S_n = \sum_i W_{i,n} [(d_i - 1)^n - X_{i,n}] / \sum_i W_{i,n} \quad (12)$$

for some set of weights $W_{i,n}$. The minimum-variance estimate will have weights inversely proportional to the variance of each term

$$W_{i,n} = 1/\text{Var} (d_i - 1)^n = 1/[E(d_i - 1)^{2n} - E^2(d_i - 1)^n] \quad (13)$$

This raises a classic problem—our estimate of, say, the second moment R_2 requires knowledge of moments up to R_4 , which in turn requires knowledge of moments up to R_8 and so on. We have chosen to close the hierarchy by calculating W_1 , W_2 and W_3 directly, and assuming that $W_4 \propto W_2^2$, $W_5 \propto W_2 W_3$ and $W_6 \propto W_3^2$. We can then form unbiased, close-to-minimum variance estimates of $R_1, \dots, R_6$, although we can only find formal uncertainties on our estimators for R_1 , R_2 and R_3 .

Starting with initial estimates of the first six central moments of $\hat{\rho}$ (setting them all to zero works well), we form weights and hence new estimates and iterate until we have a stable solution. At each step we renormalize the selection function to ensure that the weighted mean S_1 is unity. In most cases, the solution converges quickly (5–10 iterations) and the higher moments are sensible. When this does not occur, we force the higher moments to satisfy the Liapounoff inequalities²⁷ to ensure that we get a sensible weighting scheme.

The equations give us variances on the estimates for R_1 , R_2 and R_3

$$\text{Var} (S_n) = 1 / \sum_i 1/\text{Var} (d_i - 1)^n \quad (14)$$

Because we have heavily oversampled our volume, however, these variances are seriously underestimated; the formula above ignores the covariances between estimates from neighbouring grid points. We can estimate these covariances as follows. Suppose we have a white-noise random field, and we have two estimates s_n and s'_n of the n th central moment of $\hat{\rho}$ made using smoothing functions w and w' with centres r apart. Then the fractional covariance between the two estimates is simply

$$\frac{\text{Cov} (s_n, s'_n)}{\text{Var} (s_n)} = \frac{\int w^n w'^n dV}{\int w^{2n} dV} = \exp \left[-\frac{n}{4} \left(\frac{r}{r_{\text{sm}}} \right)^2 \right] \quad (15)$$

for our gaussian smoothing functions. We find this sum for all pairs of lattice points, and hence find how much the variances have been underestimated. We find that we have to scale up the variances for the first, second and third moments by 43.03, 15.74 and 8.57, respectively. The first of these tells us by how much

we have oversampled the volume, and the other two factors are included in the uncertainties quoted in Table 2.

There is further uncertainty in our answers due to statistical error in the luminosity function and evolution parameters. The 1σ uncertainty in the luminosity function takes the form of a three-dimensional ellipsoid in parameter space, and is independent of the uncertainty in the evolution³; there are then four independent axes along which we can vary the parameters. The error in the final moments can then be found by calculating solutions with parameters varied by $\pm 1\sigma$ along each axis, and adding these errors in quadrature to the variances found above.

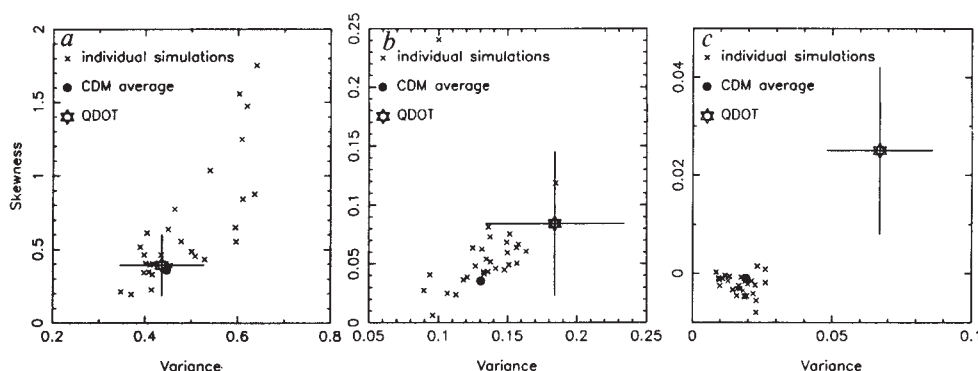
Statistics of the density field

We have run the same calculation on three-dimensional Poisson simulations, and as expected we get negligibly small underlying moments in this case. We have also run the calculation on a series of 30 standard cold dark matter (CDM) simulations of the QDOT survey, generated by methods described elsewhere^{27,28}. The simulations were made assuming a Einstein-deSitter universe with Hubble constant $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$, a bias parameter of $b_{\text{IRAS}} = 1.5$ and with small-scale correlation function matching optical galaxies. The same selection function as determined for the QDOT survey was used, and they were randomly sampled and smoothed in an identical manner.

The results for the QDOT survey and the average for the CDM simulations are shown in Table 2. The uncertainty in the results due to uncertainty in the luminosity function and evolution parameters is very much smaller than the statistical error, and has been included in the results below. In Fig. 4, we show the variance and skewness for each simulation, together with the CDM average and the QDOT result and its 1σ uncertainty. There are two main results from this analysis. First, the QDOT survey contains more power on large scales than the CDM simulations. As Fig. 4 shows, for $r_{\text{sm}} = 5h^{-1} \text{ Mpc}$ the variance and skewness of the QDOT survey are in excellent agreement with the CDM prediction, for $r_{\text{sm}} = 10h^{-1} \text{ Mpc}$ the agreement is marginal, and for $r_{\text{sm}} = 20h^{-1} \text{ Mpc}$ none of the 30 simulations comes close to the QDOT result. We can then rule out the standard CDM model to at least the 97% confidence limit. Second, the present-day galaxy density field is non-gaussian: the QDOT galaxy density field is skewed to high densities even for the largest smoothing scale, although this result is only significant at the 1.5σ level for $20h^{-1} \text{ Mpc}$ smoothing scale. This does not imply that the initial fluctuations were non-gaussian, or that there is any topological asymmetry between over- and underdense regions.

For comparison with results from different methods, we note that for a white-noise random field, the variance in gaussian spheres of radius $rh^{-1} \text{ Mpc}$ is the same as that in top-hat spheres of radius $2.2rh^{-1} \text{ Mpc}$ or cubes of side $3.5rh^{-1} \text{ Mpc}$, and the same is true for a gaussian random field with correlation function slope of -1.8 , it is equivalent to $2.2rh^{-1} \text{ Mpc}$ and $3.5rh^{-1} \text{ Mpc}$, respectively. Our answers are then consistent with the counts-in-

FIG. 4 Scatter plots of the variance and skewness of 30 individual CDM simulations, the CDM prediction and the QDOT results for *a*, $r_{\text{sm}} = 5h^{-1} \text{ Mpc}$, *b*, $r_{\text{sm}} = 10h^{-1} \text{ Mpc}$ and *c*, $r_{\text{sm}} = 20h^{-1} \text{ Mpc}$. The error bars for the QDOT results are the 1σ uncertainties.



cells analysis⁵.

The method also gives an uncertainty for the mean density of IRAS galaxies. The result for $r_{\text{sm}} = 20h^{-1}$ Mpc is 4.6%, and although this is an underestimate if there are significant correlations on scales greater than our smoothing scale, it suggests that this survey is approaching a fair sample of the Universe.

Conclusions

We have used the QDOT survey of IRAS galaxies to generate

density-field maps of the local Universe. Many of the features seen were previously known, but many new superclusters and voids are seen in poorly surveyed parts of the sky. The root-mean-square density variation falls off with smoothing scale less rapidly than predicted by the standard cold dark matter theory of galaxy formation. We have detected skewness in the density field, even for smoothing radius $20h^{-1}$ Mpc, and the assumption of gaussian density fluctuations at the present epoch on these scales is no longer secure. □

Received 18 September; accepted 30 November 1990.

1. Kaiser, N. *Mon. Not. R. astr. Soc.* **219**, 785–790 (1990).
2. Rowan-Robinson, R., Saunders, W., Lawrence, A. & Leech, K. L. *Mon. Not. R. astr. Soc.* (submitted).
3. Saunders, W. *et al. Mon. Not. R. astr. Soc.* **242**, 318–337 (1990).
4. Rowan-Robinson, M., *et al. Mon. Not. R. astr. Soc.* **247**, 1–18 (1990).
5. Efstathiou, G. *et al. Mon. Not. R. astr. Soc.* **247**, 10p–14p (1990).
6. Strauss, M. & Davis, M. in *Proc. 3rd IRAS Conf., Cornets to Cosmology* (ed. Lawrence, A.) 361–369 (Springer, Berlin, 1988).
7. Shapley, H. & Ames, A. *Ann. Harvard College Obs.* **88**(2) (1932).
8. Zwicky, F. *et al. Catalog. of Galaxies and Clusters of Galaxies* (California Institute of Technology, 1961–1968).
9. Shane, C. D. & Wirtanen, C. A. *Publ. Lick Obs.* **22**, 1 (1967).
10. Davis, M., Huchra, J., Latham, D. W. & Tonry, J. *Astrophys. J.* **253**, 423–445 (1982).
11. da Costa, L. N. *et al. Astrophys. J.* **327**, 544–560 (1989).
12. Geller, M. J. & Huchra, J. P. *Science* **246**, 897–903 (1990).
13. Loveday, J. D. thesis, Univ. Cambridge (1989).
14. Giovanelli, R. & Haynes, M. P. in *Large Scale Structures of the Universe, IAU Colloq. 130* (eds Audouze, J., Pelletan, M.-C. & Szalay, A.) 113–120 (Kluwer, 1988).
15. Tully, R. B. *Astrophys. J.* **303**, 25–38 (1986).
16. Tully, R. B. *Astrophys. J.* **323**, 1–18 (1987).
17. Bahcall, N. A. & Soneira, R. M. *Astrophys. J.* **277**, 27–37 (1984).

18. Peebles, P. J. E. *The Large Scale Structure of the Universe* (Princeton University Press, 1980).
19. Hacking, P. B., Condon, J. J. & Houck, J. R. *Astrophys. J.* **316**, L15–L19 (1987).
20. Lonsdale, C. J., Hacking, P. B., Conrow, T. P. & Rowan-Robinson, M. *Astrophys. J.* **358**, 60–80 (1990).
21. Rowan-Robinson, M. *Space Sci. Rev.* **48**, 1–71 (1988).
22. Kaiser, N. *Mon. Not. R. astr. Soc.* **227**, 1–21 (1987).
23. Lahav, O., Rowan-Robinson, M. & Lynden-Bell, D. *Mon. Not. R. astr. Soc.* **234**, 677–701 (1988).
24. Kaiser, N. & Lahav, O. *Mon. Not. R. astr. Soc.* **237**, 129–162 (1989).
25. Cox, A. H. & Miller, R. J. *The Theory of Stochastic Processes* (1974).
26. Kendall, M. G. & Stuart, A. *The Advanced Theory of Statistics* Vol. 2, 4th edn (Griffin, 1979).
27. White, S. D. M., Frenk, C. S., Davis, M. & Efstathiou, G. *Astrophys. J.* **313**, 505–516 (1987).
28. Frenk, C. S., White, S. D. M., Efstathiou, G. & Davis, M. *Astrophys. J.* **351**, 10–21 (1990).
29. Burstein, D. *et al. in Large Scale Structure and Motions in the Universe* (eds Mezzetti, M., Guiric, G. T., Mardrossian, F. & Ramella, M.) (Kluwer, 1989).
30. Fairall, A. P. & Jones, A. *Publ. Dept. Astr. Cape Town* **10** (1988).
31. Scaramella, R., Balesi-Pillastrini, G., Chincarini, G., Vettolani, G. & Zamorani, G. *Nature* **338**, 562–564 (1989).

ACKNOWLEDGEMENTS. We thank the staff at the AAT, and also those at the INT and WHT, operated by the Royal Greenwich Observatory at the Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias. We thank J. Huchra, M. Davis, M. Strauss, N. da Costa and R. Wolstencroft for supplying us with redshifts before publication. W.S. and A.L. acknowledge SERC postdoctoral and advanced fellowships, respectively.

A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome

Carolyn J. Brown, Andrea Ballabio*, James L. Rupert, Ronald G. Lafreniere, Markus Grompe*, Rossana Tonlorenzi* & Huntington F. Willard†

Department of Genetics, Stanford University School of Medicine, Stanford, California 94305, USA

* Institute for Molecular Genetics, Baylor College of Medicine, Houston, Texas 77030, USA

X-chromosome inactivation results in the *cis*-limited dosage compensation of genes on one of the pair of X chromosomes in mammalian females. Although most X-linked genes are believed to be subject to inactivation, several are known to be expressed from both active and inactive X chromosomes. Here we describe an X-linked gene with a novel expression pattern—transcripts are detected only from the inactive X chromosome (X_i) and not from the active X chromosome (X_a). This gene, called *XIST* (for X_i -specific transcripts), is a candidate for a gene either involved in or uniquely influenced by the process of X inactivation.

X-CHROMOSOME inactivation is a unique developmental regulatory mechanism, affecting the expression of genes on an entire chromosome in a *cis*-limited fashion—in the case of the human X chromosome, involving over 150 million base pairs (bp) of DNA and several thousand genes. X inactivation results in

dosage equivalence between females who have two X chromosomes and males who have one X chromosome, by randomly inactivating one of the X chromosomes in females^{1–3}. The mechanism of X inactivation remains unknown, although genetic evidence strongly suggests the existence of a specific locus required in *cis* for X inactivation to occur; this is the so-called X inactivation centre^{4,5}.

Many X-linked genes have been shown to be subject to X inactivation in both man and mouse on the basis of their mosaic or clonal expression^{3,6–8}. Although only a few genes have been analysed for transcription^{9–13}, the basis for X inactivation is generally thought to be transcriptional. A growing number of genes have been described that escape X inactivation, both in or adjacent to the pairing segment with the Y chromosome^{14,15}, as well as more proximally on the short arm and on the long arm of the human X chromosome^{9,10,16,17}. All of these genes are well expressed from both X_a and X_i chromosomes, although in some cases to a lesser extent from X_i .

To gain an insight into the molecular basis of X inactivation, we have searched for additional genes expressed from the X_i , particularly among those that map near the interval of the X inactivation centre (*XIC*) on the human X chromosome. We report here the isolation and characterization of a novel gene, *XIST*, expressed from X_i but not from X_a chromosomes, which by virtue of its localization on the X to the same interval as

† To whom correspondence should be addressed.

XIC, represents a candidate for a gene either involved in or strongly influenced by X inactivation.

Identification of X_i -specific transcripts

A 750-bp cDNA clone (8A11) was one of six clones previously obtained from a female placental complementary DNA library as part of a study to isolate cDNAs for human steroid sulphatase (ref. 18). This clone was used as a probe for the isolation of two related cDNA clones from the same library (see Fig. 1). On the

basis of the localization of the gene encoding these cDNAs to the region of the X inactivation centre in Xq13 (ref. 19), we examined its expression by RNA slot blots, by northern blot analysis and by amplification of reverse-transcribed RNA using the polymerase chain reaction (RT-PCR). As shown in Fig. 1*a*, *XIST* cDNA probes hybridized to RNA prepared from female samples or from somatic cell hybrids containing an inactive human X chromosome, but not to RNA from males or from hybrids containing only an active human X chromosome. On

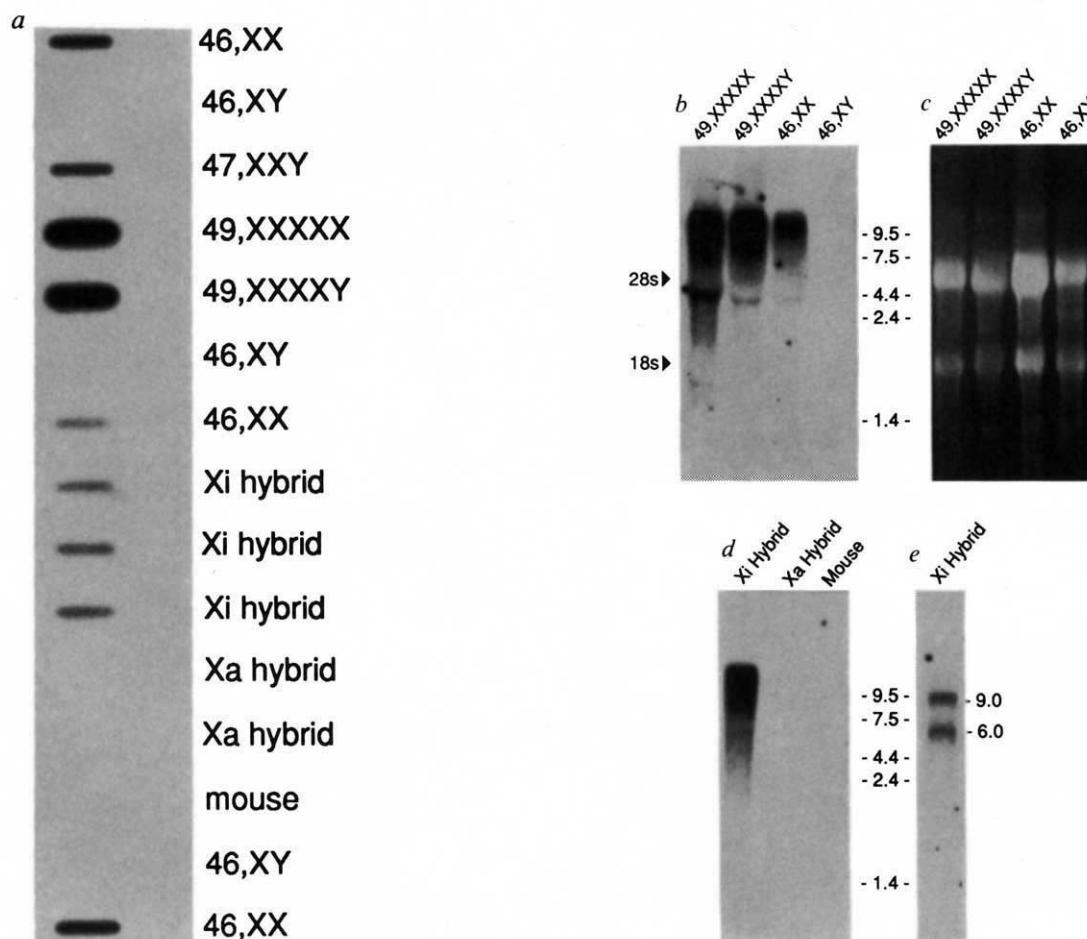


FIG. 1 Expression of the *XIST* gene in males, females, and somatic cell hybrids. *a*, Slot blot of total cellular RNA isolated from human lymphoblastoid cell lines or mouse-human somatic cell hybrids retaining either the active or inactive human X chromosome, hybridized with the 14A *XIST* cDNA probe. The probe hybridizes only to RNA samples from cell lines which contain an X_i . *b*, Northern blot of total cellular RNA from male and female cell lines, hybridized with the 14A cDNA probe. The probe detects a large, heterogeneous message only in cell lines which contain an X_i . Apparent bands below the position of 28S and 18S rRNA are created by the huge abundance of rRNA present. *c*, Ethidium bromide-stained RNA before northern transfer for the filter in *b*. *d*, Northern blot of poly(A)⁺ RNA from human/mouse hybrids retaining either the X_a or the X_i and a mouse cell line control. The 14A *XIST* probe hybridizes to transcripts only in the hybrid with an inactive X chromosome. *e*, The X_i hybrid lane of *d* stripped and reprobed with a control cDNA probe (SB1.8) identifying a 9- and a 6-kb mRNA (ref. 17) to show that the mRNA is intact.

METHODS. Total cellular RNA was prepared using the guanidinium thiocyanate method³⁷. The male (Y24.1, GM4391, GM7308) and female (Y24.2, HSC192, GM7011) lymphoblast cell lines (panel *a*, top to bottom in each case) were from karyotypically normal individuals. The X chromosome aneuploid lines were: 47, XXY, D64.0; 49, XXXXY, GM1202; 49, XXXXX, GM6061B. The X_a hybrids were t60-12 and AHA-11aB1, both of which contain the active X

chromosome as the only human chromosome³⁸. The X_i hybrids were t11-4Aaz5; t48-1a-1Daz4A and t86B1maz1B (ref. 38). 10 µg of total cellular RNA was denatured in formaldehyde/formamide and applied to a nitrocellulose membrane without vacuum for the slot blots. As controls, the RNA samples were extensively digested with RNase-free DNase I before applying to slot blots; the results were equivalent to those shown. To ensure equal RNA loading, the blot was stripped and reprobed with a cDNA for the human *PGK1* gene; roughly equal signals were produced from each of the human RNA slots. For the northern blots, 20 µg total RNA or 5 µg poly(A)⁺ RNA (purified by one or two passages through oligo(dT)) was electrophoresed on a 1.5% agarose gel in the presence of formaldehyde and transferred by capillarity to a nitrocellulose membrane. The blots were hybridized using standard methods¹³ with ³²P-labelled 14A insert (*a*, *b*, *d*) or SB1.8 (ref. 17) (*e*), and washed at 42 °C with 0.1% SDS; 0.5 × SSC. The 8A11 cDNA clone was one of six clones identified by immunoscreening of a placental λ gt11 cDNA expression library, using anti-steroid sulphatase antibodies¹⁸. The basis for the apparent immunoreactivity of this clone has not been investigated further; the clone has no significant DNA homology with the steroid sulphatase cDNAs reported¹⁸. The cDNA clones 14A and 7A were isolated by filter hybridization from the same library, using 8A11 as probe. The 14A cDNA clone had the longest insert (see Fig. 3) and was used as probe for all hybridizations.

northern blots, the same pattern of female (46,XX)- and X_i -specific hybridization was observed with a signal indicating a series of multiple transcripts showing extensive size and/or conformational heterogeneity, most >10 kb (Fig. 1b,d). Although in some experiments the hybridization pattern was suggestive of discrete bands, in most cases the signal was continuous, even when smaller cDNA fragments were used as probes (data not shown). Control experiments verified the integrity of the RNA samples tested (Fig. 1e), and because no hybridization was detected with the *XIST* cDNA probes in male RNA samples, hybridization artefacts could be discounted. In

both slot blots and northern blots, the intensity of hybridization appeared to increase with the number of inactive X chromosomes present (Fig. 1a,b), although more extensive quantitation is required to confirm this. The northern blot data therefore suggest many heterogeneous *XIST* transcripts expressed from X_i but not from X_a chromosomes.

Because of the heterogeneity of transcripts detected with the cDNA probe, oligonucleotide primers were used for analysis by RT-PCR. PCR product was only observed in those reverse-transcribed RNA samples isolated from sources which contained an X_i . Karyotypically normal females were positive for the

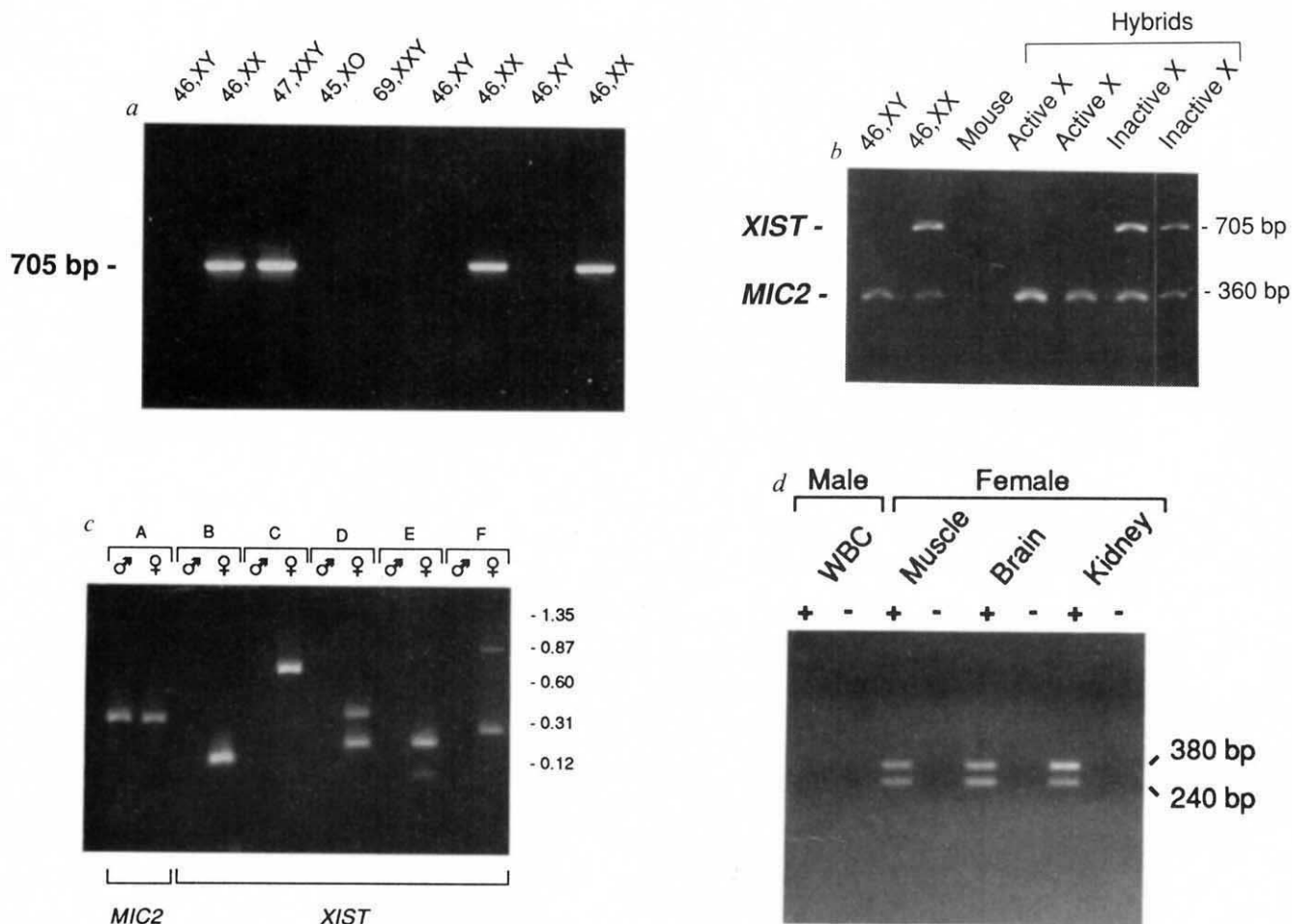


FIG. 2 RT-PCR analysis of *XIST* expression. *a*, RT-PCR of reverse-transcribed RNA (RT-RNA) from lymphoblastoid cell lines from karyotypically normal males and females, and individuals with X chromosome aneuploidies. The primers (sequence below, see also Fig. 3) amplified a 705-bp product only in RT-RNA from individuals with an X_i . *b*, RT-PCR of somatic cell hybrid lines retaining either X_a or X_i . As in *a*, the *XIST* primers amplified a 705-bp product only in those cell lines which contained an X_i . *MIC2* primers were included as a duplexed control for the presence of amplifiable cDNA, as *MIC2* is expressed from both the X_a and the X_i . The 360-bp human-specific *MIC2* product was observed in all hybrid lanes. *c*, RT-RNA from a male and female lymphoblast cell line was used in a PCR assay with five different *XIST* primer pairs (lanes B to F), and the *MIC2* primers as a control (lanes A). The primers are diagrammed in Fig. 3, and their sequences are given below. The 360-bp *MIC2* product was synthesized from both male and female RT-RNA, but the *XIST* products were observed only from the female RT-RNA. More than one product (only from the female RT-RNA) was observed for primer pairs D and F. Similar results are observed with hybrids containing the X_i (data not shown). The numbers to the left of the figure represent the size (in kb) of the *HaeIII*-digested ϕ X174 markers. *d*, RT-PCR analysis of *XIST* expression in samples from different tissues. RNA was prepared from male white blood cells (WBC) and from female brain, kidney, and muscle and was assayed both with (+) and without (-) reverse transcriptase.

METHODS. RNA (5 μ g), isolated as described in Fig. 1, was reverse-transcribed with MMLV reverse transcriptase using random hexamer primers^{13,39}. One-hundredth of the cDNA obtained was used for a PCR reaction with 1 μ mol primers, for 30 cycles (94 °C, 1 min; 55 °C, 1 min; 72 °C, 4 min) with Promega or Cetus *Taq* polymerase and buffer. *d*, RNA was prepared from tissues obtained from an 18-week female fetus. Primers are listed 5'-3' with the arrow designating orientation on the gene. *a*, Primers were: 1 $\rightarrow$: GAAGTCTCAAGGCTTGAGTTAGAAG; and 3 $\leftarrow$: GCCAGGCTCTAGAGAAAATGT. *b*, *XIST* primers were as in *a* and the *MIC2* primers were as described previously¹³. PCR amplification of RNA samples which have not been reverse-transcribed does not yield the 705-bp *XIST* product or the 360-bp *MIC2* product (data not shown; see also *d*). *c*, Primer pairs were: A, *MIC2* primers, as described¹³; B, 1 $\rightarrow$ (GAAGTCTCAAGGCTTGAGTTAGAAG) and 2 $\leftarrow$ (AACCCAGGAGATAGGTAGATCCATC) to give a 180-bp product; C, as in *a* to give a 705-bp product; D, 3 $\rightarrow$ (GCCAGGCTCTAGAGAAAATGT) and 5 $\leftarrow$ (ACAGACGTATTTTCGTCTAA) which give a 240- and a 380-bp product; E, 3 $\rightarrow$ (GCCAGGCTCTAGAGAAAATGT) and 4 $\leftarrow$ (TGGCTCAAGTGTAGGTGGTT) which give a 260-bp product; F, 1 $\rightarrow$ (GAAGTCTCAAGGCTTGAGTTAGAAG) and 5 $\leftarrow$ (ACAGACGTATTTTCGTCTAA) which give a 300-, a 900- and a 1,040-bp product. Location of primers in the *XIST* gene are indicated in Fig. 3. *d*, Primer pair D was used.

TABLE 1 Summary of *XIST* expression

Cell-line	Karyotype	RT-PCR	Expression Slot blot	Northern
Chromosomally normal				
7 females	46, XX	+(7)	+(6)	+(6)
10 males	46, XY	-(10)	-(6)	-(6)
Chromosomally aberrant				
GM6061B	49, XXXXX	+	+	+
GM1202	49, XXXXY	+	+	+
106	45, XO	-	-	-
107	69, XXY	-	-	-
D64.0	47, XXY	+	+	+
GM10074 (ref. 19)	47, Y, t(X; 14)	+	+	+
	+der(14)	+	ND	ND
A.G. (ref. 19)	45, X/46, X, idic(Xp)	+	ND	ND
Somatic cell hybrids				
t60-12	Active X	-	-	-
AHA11aB1	Active X	-	-	-
t11-4Aaz5	Inactive X	+	+	+
t48-1a-1Daz4a	Inactive X	+	+	+
t81-az1D	Inactive X	+	+	+
t86-B1maz1b-3a	Inactive X	+	+	ND
LT23-1E2Buv5C126	Inactive X	+	ND	ND

The female RNAs analysed were from six lymphoblastoid cell lines and five different tissues from a single female fetus (see text and Fig. 2d). The male RNAs were isolated from lymphoblastoid cell lines (7 males), liver (2 males) and white blood cells (1 male). The '+' and '-' indicate whether expression was detectable for each assay (see Figs 1 and 2). 'ND' indicates that the assay was not performed for the given cell line. The numbers in parentheses after the '+' or '-' indicate the number of different female or male RNAs for which the specified assay was performed.

METHODS. RNA was isolated from the specified cells and analysed as described for Figs 1 and 2. All somatic cell hybrids were derived from female human cells and are described in Fig. 1, except for the X_i hybrids LT23-1E2Buv5C126, which was formed by fusion of mouse Ltk⁻ cells with a 49, XXXXX human cell line, and t81-az1D, which was from a fusion of mouse tsA1S9 cells and the human cell line 81 (ref. 19). The four X_i hybrids t11-4Aaz5, t48-1a-1Daz4a, t81-az1D, and t86-B1maz1b-3a were under selection¹⁶ for the X_i ; these hybrids retain other human chromosomes, although no chromosome except the inactive X is common to all four lines. To eliminate any effect of other human chromosomes or any influence caused by selection, a fifth independent X_i -containing somatic cell hybrid was examined. This hybrid line, LT23-1E2Buv5, was not under selection for the X_i , and molecular cytogenetic analysis has identified that the only human material in this cell line is the X_i chromosome, present in about one-third of cells examined (V. E. Powers, data not shown).

705-bp product, whereas karyotypically normal males were negative (Fig. 2a). To determine if expression was sex-specific, in addition to X_i -specific, RNAs from a female without an inactive X (45, XO) and from males with one or more inactive X's (47, XXY; see also the 49, XXXXY sample in Fig. 1a) were examined. As shown in Fig. 2a, RT-PCR of RNA from the 45, XO female did not generate a product, but the RNA of the 47, XXY male did. RNA from a triploid cell line (69, XXY), in which both X chromosomes are active^{20,21}, was also examined. Despite the presence of two X chromosomes, no PCR product was detected. In total, samples from seven normal females, ten normal males, and seven chromosomally aberrant cell lines were analysed for *XIST* expression by PCR, northern blot and/or RNA slot blot (Table 1). In all cases, *XIST* expression was only observed in the presence of an X_i .

Additional evidence for X_i -limited expression came from the analysis of somatic-cell hybrids containing either X_i or X_a chromosomes (Fig. 2b). Hybrids with an X_a did not express *XIST*, whereas five independent hybrids with an X_i did (Table 1). As a control, all hybrids were shown to produce an RT-PCR product corresponding to the X-linked *MIC2* gene, consistent

with previous demonstrations that *MIC2* escapes inactivation^{13,15}. Taken together, the data in Figs 1 and 2 and Table 1 support the conclusion that expression of the *XIST* transcripts is absolutely dependent on the presence of an inactive X chromosome.

XIST gene structure and alternative transcripts

To examine expression of the *XIST* gene in more detail, several different oligonucleotide primers along the gene were synthesized for RT-PCR analysis. For each primer pair tested (see Fig. 3), a product was observed only in female or X_i -containing hybrid RNA samples, and never in male or X_a -containing hybrid RNA samples (Fig. 2c). A number of the *XIST* primer pairs gave more than one RT-PCR product, suggesting that alternatively spliced RNAs are present. Multiple transcripts were observed in all female tissues tested, including heart, muscle, brain, kidney, liver, fibroblasts and lymphoblasts (Fig. 2d and data not shown).

To identify portions of the *XIST* gene, two overlapping genomic clones were isolated which contained over 20 kb of genomic DNA corresponding to the *XIST* cDNAs (see Fig. 3). A genomic restriction map was derived from both Southern blot analysis of genomic DNA and mapping of the various clones. Both the intensity and pattern of hybridization on Southern blots was identical in normal males and in hybrids containing only a single X, at both normal and reduced stringency; thus all evidence is consistent with there being only a single copy of the *XIST* locus located on the X chromosome. Transcribed portions of the *XIST* gene were identified in cDNA clones isolated from two different cDNA libraries (see Fig. 3) and in various RT-PCR products isolated from both female cell lines and X_i -containing hybrid lines.

Several alternatively spliced RNA products have been identified and are diagrammed in Figs 3 and 4. The complete nucleotide sequence of the original cDNAs was determined, and all novel RT-PCR products not represented in the original cDNA clones (Fig. 3) were also sequenced. The sequences have been compared with the genomic *XIST* sequence across each putative splice junction (see Fig. 4). All putative donor and acceptor splice sites agree completely with consensus splice site sequences²².

By using the RACE extension protocol²³ to identify 3' poly(A) tails on RT-PCR products, a number of differentially spliced products were observed. The primary RACE product ended at the poly(A) site marked A1 on the map in Fig. 4. A second, less common product predicts the poly(A) site marked A2. The orientation of the RACE products, as well as the nature of the splice junctions, indicates the gene orientation shown in the figures. In total, at least five different splice products have been observed within this ~7-kb stretch of genomic DNA. This degree of alternative splicing, if representative of the entire gene, would explain the heterogeneous *XIST* RNA signal observed on northern blots (Fig. 1). Knowledge of the full extent of alternative splicing awaits isolation and analysis of the entire gene sequence, which the northern blot data suggest is likely to be at least 10 kb long.

Over 2.2 kb of transcribed *XIST* sequence has been determined (Fig. 5). Multiple stop codons were detected in all six reading frames and there were no sustained open reading frames >~300 bp in length for any of the predicted splice products observed to date. Furthermore, none of the sequences scored significantly in an algorithm designed to detect coding sequences²⁴. It is likely, therefore, that the isolated portion of the *XIST* gene corresponds largely to untranslated material. Given the absence of an extended open reading frame, the fact that the original cDNA clone was identified by immunoscreening with anti-steroid sulphatase antibodies¹⁸ remains unexplained. The entire sequence has been deposited in the GenBank/EMBL Data Library (accession no. X56196-X56199); no sequences in GenBank (August 1990 version), including steroid sulphatase

or any other X-linked gene for which sequence is available, show significant homology with any portion of the *XIST* sequence.

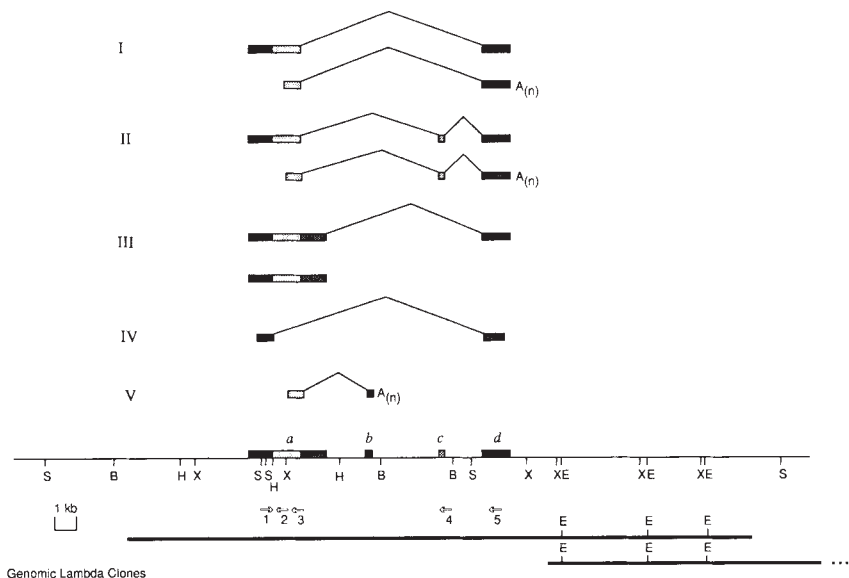
Discussion

The X inactivation centre (*XIC*) is a locus required in *cis* for an X chromosome to be inactivated¹⁻⁵. The Xq13 region of the

human X chromosome, which contains the *XIC*^{25,26}, is also the site of Barr-body condensation²⁷ and the site of a cytological bend on X_i chromosomes²⁸. The region is homologous to a region on the mouse X chromosome, also required for X inactivation^{29,30} and containing the *Xce* locus, alleles at which affect the choice of X chromosome to be inactivated^{31,32}. The close association of these various inactivation-related

FIG. 3 Genomic and cDNA map of the *XIST* gene.

The schematic drawings represent different transcribed products (I-V) from the *XIST* gene. The schematic drawings represent different transcribed products (I-V) from the *XIST* gene as recognized by analysis of cDNA clones or RT-PCR products. All products were X_i -specific. Product I is the 14A cDNA clone, and is consistent with the RT-PCR products from primers 1→ and 2← (primer pair B in Fig. 2c); 1→ and 3← (primer pair C); the lower product of 3→ and 5← (primer pair D); and the 900-bp product of 1→ and 5← (primer pair F). Product II is colinear with product I except for the addition of a 140-bp exon. This product has been identified as cDNA clone 7A (see Fig. 1) and is consistent with the RT-PCR products listed above, except that with primer pair D this product is the upper, 380-bp product, and with primer pair F it is the upper 1,040-bp product. Primer 4 is located within this additional 140-bp exon, and the observed 260-bp product from primers 3→ and 4← (primer pair E of Fig. 2c) is consistent with this product. Use of the RACE 3' extension protocol with the 3→ primer yielded a number of products, of which the two major products are consistent with the splices of products I and II, with a 3' extension including a poly(A) tail. These products are diagrammed below products I and II. Product III was isolated as a cDNA clone (pJR-c1) from a Stratagene human female fetal brain library (cat. no. 936206) by filter hybridization using 14A as probe. The 8A11 cDNA clone is identical to pJR-c1 up to the position of the splice, at which point 8A11 ends. Product III is also consistent with the RT-PCR products from primer pairs B and C. Product IV has only been observed as the RT-PCR product of 300 bp from the 1→ and 5← primers (primer pair F of Fig. 2c). Product V has been identified as a minor product of the 3' RACE extension from primer 3→. Below these various splice products is a diagram of the *XIST* genomic DNA. This map is derived from Southern blotting of genomic DNA



and is consistent with the map of the cDNAs described above, and the two genomic λ phage diagrammed at the bottom of the figure. The *a*, *b*, *c* and *d* above the transcribed portions of the *XIST* gene correspond to the sequenced segments (Fig. 5). The genomic phage clones were isolated by plaque hybridization with the 14A clone of the Los Alamos flow-sorted human X chromosome library LAOXN01 (obtained from the NIH Probe and Library Repository at the American Type Culture Collection, Rockville). The primer locations are represented by open arrows (not to scale). The arrow points in the direction in which the primer is 5' to 3'. S, *Stu*I; B, *Bgl*II; H, *Hind*III; X, *Xba*I; E, *Eco*RI.

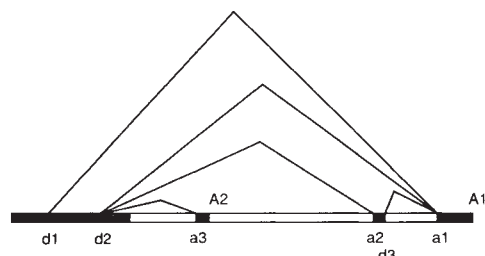


FIG. 4 Alternatively spliced cDNA products, and sequence of the splice junctions and poly(A) addition sites. *a*, The different splice junctions diagrammed in Fig. 3 are summarized in this drawing of the *XIST* gene, and the splice junctions are designated as donors (d1-d3) or acceptors (a1-a3). The alternate polyadenylation sites as determined by RACE 3' extensions are marked A1 and A2. A1 is a more abundant product than A2. *b*, The sequence of the different splice junctions. The '/' represents the splice site as determined by comparing the genomic and cDNA sequences. cDNA sequences are indicated in upper case. Genomic sequences not represented in cDNAs are indicated in lower case. All splice junctions agree with the consensus splice sequence²² at the invariant bases (underlined), and generally agree at most of the favoured bases. The product designation refers to the products diagrammed in Fig. 3. *c*, Sequence of genomic DNA and cDNA before the polyadenylation sites A1 and A2. For both sites there is a potential polyadenylation signal (underlined) within 20 bp of the poly(A) tail.

Splice donor	Splice acceptor	Product
d1 - AAG/gtatct	a1 - ctttcacag/TIT	IV
d2 - TAG/gtgagc	a1 - ctttcacag/TIT	I
d2 - TAG/gtgagc	a2 - ttttctag/TCC	II
d3 - AAG/gtttgt	a1 - ctttcacag/TIT	II
d2 - TAG/gtgagc	a3 - ttttccag/GTA	V
AAG/gtaagt	Pg ₍₆₎ ncag/N	Consensus

A1	genomic DNA	CTTCTGTAT GAATAAACT TTCTGTAGC TAGGTATTGT CTCTAC
	RACE cDNA	CTTCTGTAT GAATAAACT TTCTGTAGC TAGAAAAA AAAAAA
A2	genomic DNA	TGCTATTGTA AATAAATA TACTGTTTG AAACTTTGT ATTTGA
	RACE cDNA	TGCTATTGTA AATAAATA TACTGTTTG AAAAAAAA AAAAAA

phenomena establishes the primary importance of this region in X inactivation. As shown in the accompanying paper¹⁹, the *XIST* gene is located in the portion of band Xq13 common to all structurally abnormal X_i chromosomes, and is therefore coincident with the *XIC* at the current level of mapping resolution.

The novel expression of *XIST*, its absolute dependence on X_i chromosomes and its mapping to the region of the *XIC* are strongly suggestive of some involvement in the X inactivation process. As the mechanism of X inactivation is not known, the role of the *XIC* and the potential involvement of *XIST* are necessarily conjectural. The involvement could be primary, in which case the *XIST* gene would be predicted to have a critical function in X inactivation, or secondary, in which case its expression might be determined as a result of other events occurring nearby which result in X inactivation; nonetheless, its X_i -specific and, among normal individuals, female-specific pattern of expression is novel among X-linked genes described to date.

X chromosome inactivation was first hypothesized by Lyon³. Much data have since accumulated supporting the theory that most X-linked genes are stably transcriptionally inactivated early in development; however, little is known either about the initi-

ation of inactivation at the *XIC* early in embryogenesis or about the spreading of inactivation in *cis* along the length of the chromosome. All X chromosomes in excess of one are inactivated, suggesting that some signal prevents one X (the eventual X_a) from responding to a process which affects all unmarked X chromosomes (the eventual X_i 's)³³. It is reasonable to hypothesize that this signal acts on the *XIC*, as the *XIC* is the only region of the X which is required in *cis* for X inactivation to occur^{19,25,26}. This signal would be responsible for blocking the marked X chromosome from responding to the developmental signal for inactivation^{29,33,34}.

If *XIST* expression has a primary role in X inactivation, then the hypothesized signal may interact directly with the *XIST* locus, thereby preventing its expression from active X chromosomes. Given the *cis*-limited nature of X inactivation³⁵, expressed products from the *XIST* gene on the X_i would be required to act in a similar *cis*-limited manner, if this gene is involved in X inactivation. One possibility may be that the *XIST* product is a *cis*-acting RNA molecule, perhaps involved structurally in formation of the heterochromatic Barr body. It is known, however, that at least some of the *XIST* RNA is polyadenylated (Fig. 1d), which would be more suggestive of a translatable mRNA. If *XIST* is a protein-encoding gene, its

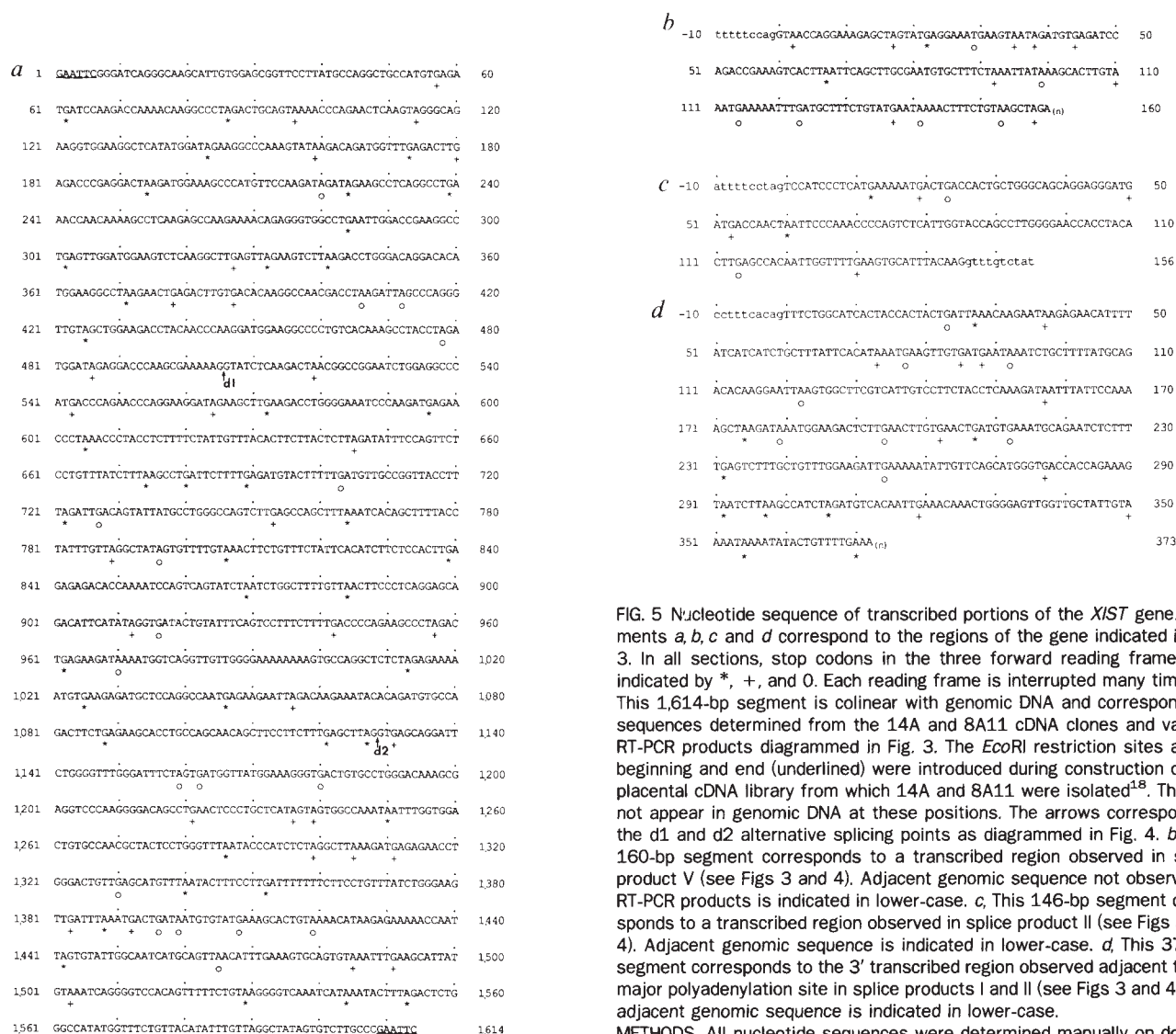


FIG. 5 Nucleotide sequence of transcribed portions of the *XIST* gene. Segments a, b, c and d correspond to the regions of the gene indicated in Fig. 3. In all sections, stop codons in the three forward reading frames are indicated by *, +, and 0. Each reading frame is interrupted many times. a, This 1,614-bp segment is colinear with genomic DNA and corresponds to sequences determined from the 14A and 8A11 cDNA clones and various RT-PCR products diagrammed in Fig. 3. The *EcoRI* restriction sites at the beginning and end (underlined) were introduced during construction of the placental cDNA library from which 14A and 8A11 were isolated¹⁸. They do not appear in genomic DNA at these positions. The arrows correspond to the d1 and d2 alternative splicing points as diagrammed in Fig. 4. b, This 160-bp segment corresponds to a transcribed region observed in splice product V (see Figs 3 and 4). Adjacent genomic sequence not observed in RT-PCR products is indicated in lower-case. c, This 146-bp segment corresponds to a transcribed region observed in splice product II (see Figs 3 and 4). Adjacent genomic sequence is indicated in lower-case. d, This 373-bp segment corresponds to the 3' transcribed region observed adjacent to the major polyadenylation site in splice products I and II (see Figs 3 and 4). The adjacent genomic sequence is indicated in lower-case.

METHODS. All nucleotide sequences were determined manually on double-stranded templates using a modification⁴⁰ of the dideoxynucleotide termination method or automatically using an Applied Biosystems AB1370A fluorescent sequencer, as described by Gibbs *et al.*⁴¹.

involvement in a *cis*-limited function is more difficult to explain. Its mRNA would have to leave the nucleus to be translated on cytoplasmic ribosomes. Upon return of the *XIST* protein to the nucleus, such a product would be required to distinguish between X_a and X_i chromosomes, including, possibly, distinguishing between different copies of the *XIC* and/or different copies of the *XIST* gene itself. Presumably, the protein-encoding capacity (or incapacity) of the *XIST* gene will become apparent once the complete gene is isolated and sequenced.

A number of models would be consistent with a primary role for *XIST* in X chromosome inactivation. First, the *XIST* product could be the molecule which acts in *cis* to cause X inactivation; second, *XIST* could be the *XIC* site which acts in conjunction with a different product to cause inactivation; or third, the *XIST* product could act upon the *XIC*, which would then produce or interact with the inactivating molecule(s). It is also possible that *XIST* is not the primary target of the developmental signal, but that an exogenous, developmentally regulated molecule interacts at the *XIC* to cause X inactivation, secondarily resulting in the expression of *XIST* from X_i chromosomes. In any case, the final result of X inactivation is that a *cis*-limited signal inactivates most genes on the X_i , although there are an increasing number of genes described on both the short and long arms of the X that escape X inactivation and thus do not receive a signal, do not respond to the signal, or cannot maintain their inactive state^{9,10,14-17}.

Examination of these and other possibilities concerning the *XIST* gene would be aided by experiments of two kinds. First, it will be critical to evaluate the expression of *XIST* in early development, at the time X inactivation occurs. Presumably, for

XIST to have a role in X inactivation its expression must be modified in some critical way at the time of X inactivation. Second, it will be important to extend these observations to other species, most particularly to mouse, which offers a much more tractable system for developmental studies. Both genomic and cDNA *XIST* sequences are conserved at the level of DNA hybridization in a number of other mammalian species (including dog, cat, rabbit and cow and, at somewhat reduced stringency, mouse) (data not shown). If *XIST* expression is important in X inactivation, one would predict that homologous mouse sequences (*Xist*) should map to the homologous location on the mouse X chromosome^{30,36}, near the *Xce* locus, and should show the same pattern of X_i -restricted expression. The isolation of the homologous mouse gene would also allow for comparison of *Xist* expression and structure in strains carrying different alleles at the *Xce* locus.

Apart from its potential primary role in the X inactivation process itself, the *XIST* gene displays two novel properties, both of which are of considerable interest in their own right. First, expression is limited to inactive X chromosomes; even if *XIST* is not directly related to the *XIC* locus, it is thus strongly influenced by X inactivation and likely to provide an important molecular marker for X inactivation and an experimental foothold in the region. Second, expression of *XIST*, at least in the portion of the gene isolated to date, is normally female-specific. Identification of a potential gene product would be significant for examining developmental differences between the sexes and for analysing the potential effect of this gene's inappropriate expression in individuals with abnormalities in sex chromosome constitution, for example in Turner syndrome (45,X) and Klinefelter syndrome (47,XXY). □

Received 16 October; accepted 13 November 1990.

1. Lyon, M. F. *Biol. Rev.* **47**, 1-35 (1972).
2. Gartler, S. M. & Riggs, A. D. *A. Rev. Genet.* **17**, 155-190 (1983).
3. Lyon, M. F. *Nature* **190**, 372-373 (1961).
4. Russell, L. B. *Science* **140**, 976-978 (1963).
5. Cattanach, B. M. *A. Rev. Genet.* **9**, 1-18 (1975).
6. Migeon, B. R., der Kaloustian, V. M., Nyhan, W. L., Young, W. J. & Childs, B. *Science* **160**, 425-427 (1968).
7. Meyer, W. J., Migeon, B. R. & Migeon, C. J. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1469-1472 (1975).
8. Davidson, R. G., Nitowsky, H. M. & Childs, B. *Proc. natn. Acad. Sci. U.S.A.* **50**, 481-484 (1963).
9. Schneider-Gadicke, A., Beer-Romero, P., Brown, L. G., Nussbaum, R. & Page, D. C. *Cell* **57**, 1247-1258 (1989).
10. Fisher, E. *et al. Cell* **63**, 1205-1218 (1990).
11. Nadon, N., Dorn, N. & DeMars, R. *Somat. Cell molec. Genet.* **14**, 541-552 (1988).
12. Graves, J. A. M. & Gartler, S. M. *Somat. Cell molec. Genet.* **12**, 275-280 (1986).
13. Brown, C. J., Flenniken, A., Williams, B. & Willard, H. F. *Nucleic Acids Res.* **18**, 4191-4195 (1990).
14. Shapiro, L. J., Mohandas, T., Weiss, R. & Romeo, G. *Science* **204**, 1224-1226 (1979).
15. Goodfellow, P., Pym, B., Mohandas, T. & Shapiro, L. J. *Am. J. hum. Genet.* **36**, 777-782 (1984).
16. Brown, C. J. & Willard, H. F. *Am. J. hum. Genet.* **46**, 273-279 (1990).
17. Brown, C. J., Barker, D. & Willard, H. F. *Am. J. hum. Genet.* **45**, A178 (1989).
18. Ballabio, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 4519-4523 (1987).
19. Brown, C. J. *et al. Nature* **349**, 82-84 (1991).
20. Migeon, B. R., Sprengle, J. A. & Do, T. T. *Cell* **18**, 637-641 (1979).
21. Willard, H. F. & Breg, W. R. *Somat. Cell Genet.* **6**, 187-188 (1980).
22. Shapiro, M. B. & Senapathy, P. *Nucleic Acids Res.* **15**, 7155-7175 (1987).
23. Frohman, M. A., Dush, M. K. & Martin, G. R. *Proc. natn. Acad. Sci. U.S.A.* **85**, 8998-9002 (1988).

24. Fickett, J. W. *Nucleic Acids Res.* **10**, 5303-531 (1982).
25. Mattei, M. G., Mattei, J. F., Vidal, I. & Giraud, F. *Hum. Genet.* **56**, 401-408 (1981).
26. Summitt, R. L., Tipton, R. E., Wilroy, R. S. Jr., Martens, P. & Phelan, J. P. *Birth Defects: Orig. Art. Ser.* **14**, 219-247 (1978).
27. Therman, E., Sarto, G. E. & Patau, K. *Chromosoma* **44**, 361-366 (1974).
28. Fiejer, W. L., Van Dyke, D. L. & Weiss, L. *Am. J. hum. Genet.* **36**, 218-226 (1984).
29. Rastan, S. *J. Embryol. exp. Morph.* **78**, 1-22 (1983).
30. Rastan, S. & Robertson, E. J. *J. Embryol. exp. Morph.* **90**, 379-388 (1985).
31. Johnston, P. G. & Cattanach, B. M. *Genet. Res.* **37**, 151-160 (1981).
32. Rastan, S. *Genet. Res.* **40**, 139-147 (1982).
33. Brown, S. W. & Chandra, H. S. *Proc. natn. Acad. Sci. U.S.A.* **70**, 195-199 (1973).
34. Riggs, A. D. *Aust. J. Zool.* **37**, 419-441 (1990).
35. Lyon, M. F. *et al. J. Embryol. exp. Morph.* **97**, 75-85 (1986).
36. Keer, J. T. *et al. Genomics* **7**, 566-572 (1990).
37. Chirgwin, J. M. *et al. Biochemistry* **18**, 5294-5299 (1979).
38. Brown, C. J. & Willard, H. F. *Am. J. hum. Genet.* **45**, 592-598 (1989).
39. Grompe, M., Muzny, D. M. & Caskey, C. T. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5888-5892 (1989).
40. Korneluk, R. G., Quan, F. & Gravel, R. A. *Gene* **40**, 317-323 (1985).
41. Gibbs, R. A., Nguyen, P.-N., McBride, L. J., Koepf, S. M. & Caskey, C. T. *Proc. natn. Acad. Sci. U.S.A.* **86**, 1919-1923 (1989).

ACKNOWLEDGEMENTS. We thank S. Cohen, G. Barsh, U. Franke, D. Botstein, D. H. Ledbetter, K. E. Davies and C. T. Caskey for discussions; F. Gonzales for the placental cDNA library; G. Persico for support; and R. Gibbs, D. Muzny and A. Civitello for assistance with DNA sequencing. M.G. is supported by the National Institutes of Health. This work was supported by the National Institutes of Health (H.F.W.).

Black-hole mergers and mass inflation in a bouncing universe

A. E. Sikkema* & W. Israel

CIAR Cosmology Program, Theoretical Physics Institute,
University of Alberta, Edmonton, Canada, T6G 2J1

THE idea that our expanding Universe was born in a 'bounce'—the re-expansion of a previously contracting universe—is an old cosmological hypothesis, and continues to resurface^{1–5} despite being overshadowed recently by hypotheses inspired by Guth's inflationary model. Here we show how recent developments in the physics of black-hole interiors force a major revision of our ideas of the final moments of a contracting universe, and remove a thermodynamic difficulty^{6,7} which had appeared to rule out any kind of bounce origin for our Universe. As the black hole formed by the collapse of a rotating star settles down, it absorbs part of the gravitational radiation emitted during the last moments of collapse. This radiation, strongly blue-shifted near the inner horizon, enormously increases the mass of the black hole's core. External observers cannot detect this mass, but it manifests itself dramatically when the black holes in a collapsing universe merge, a few minutes before the 'big crunch'. The mass of a rebounding universe is enormously inflated, and its specific entropy correspondingly reduced. This allows the expansion to begin from a state of relatively low disorder.

From thermodynamic considerations, Tolman⁸ showed that the amplitude and period of a closed universe must grow at each bounce as a result of the steady increase of entropy. As the mechanical energy of expansion and contraction is gradually dissipated into heat (which is a form of inertial mass-energy) the universe comes down a little harder at each crunch, then bounces to a slightly larger radius at the next cycle. (This growth of mechanical and thermal energy is fed by the negative potential energy of the gravitational field; the total energy of a closed universe is and remains zero.) On this basis, it has been estimated^{2,9,10} that, if we live in an oscillatory universe, we can be at most 100 cycles away from the first cycle that lasted long enough to produce a generation of stars.

But there is an argument, in essence due to Penrose^{6,7}, which disallows even a single bounce. Our Universe, or any roughly similar progenitor, is expected to harbour black holes with masses ranging from stellar to galactic. During contraction, these black holes must merge, and just before the bounce, the entropy of a collapsing closed universe of mass M should therefore have grown to $s_H \approx M^2$, the area and entropy of the final all-encompassing black hole. (We use Planck units ($c = G = \hbar = k = 1$), in which the units of time, mass and temperature are roughly 10^{-43} s, 10^{-8} kg and 10^{32} K. In Planck units, the accepted age of the Universe as well as the mass and radius of the portion presently observable are of the order of 10^{60} .)

But s_H is far larger than the radiative entropy s_{rad} presently observed. This resides chiefly in the cosmic microwave background and, at 10^8 photons per baryon, amounts to $s_{\text{rad}} \approx 10^{27} M$. To beget our Universe, the entropy of a collapsing progenitor would have to deflate by a factor $s_H/s_{\text{rad}} \approx 10^{-27} M > 10^{33}$ at the bounce. In whatever manner our Universe was formed "the Creator seems to have made use of only one part in $e^{s_H/e^{s_{\text{rad}}}} > 10^{10^{120}}$ of the available chaos"⁷.

But is this reasoning correct? The final entropy s_H of a contracting universe, calculated in this manner, implies a local entropy density $S_H \approx M\rho$ that enormously exceeds the Planck value a few Planck times before the crunch, because the energy density $\rho \approx t^{-2}$ is then of the order of unity. On the other hand, general arguments based on the uncertainty principle^{11,12} suggest that planckian scales represent upper limits to attainable values

of quantities such as particle energies, temperature and entropy density. As our derivation of s_H relies on the Bekenstein-Hawking expression for the entropy of a black hole, which has at least equally solid quantum-mechanical foundations, we are confronted with a paradox.

To resolve the paradox we use what has recently been learned about black-hole interiors^{13,14}. The interior of a black hole is often considered a subject for metaphysics, but when black holes merge in the final stages of a crunch, internal processes come into play.

A generic (rotating) gravitational collapse generates a decaying wake of gravitational waves, as the external field and the outer boundary (event horizon) of the resulting black hole settle, like a quivering soap bubble, asymptotically towards the Kerr geometry, the simplest configuration compatible with the externally measured mass and angular momentum.

The Kerr geometry admits two stationary photon surfaces^{14–16}. The outer or event horizon is associated with infinite retarded time, that is, with the last photons that can still escape the collapsing object; these take arbitrarily long times to reach an external observer. The inner or Cauchy horizon is characterized by infinite advanced time. It is associated with the last photons, propagating inward from infinity, that can still be received by an observer who has fallen into the hole. In the brief moment before this observer plunges through the Cauchy horizon, the entire future history of the outer Universe will be flashed in fast motion before his eyes!

This infinite time compression is attended by an arbitrarily large blueshift of infalling radiation near the Cauchy horizon. The tail of gravitational waves left by the collapse, part of which is back-scattered or radiated directly into the hole, and which decays as an inverse power of advanced time¹⁷, becomes blueshifted (exponentially in advanced time) as it nears the Cauchy horizon. The effects on the gravitational field of the core (the region inside the Cauchy horizon) are cataclysmic. With a hundred or so horizon-crossing times, $\sim 10^2 m$ (in terms of external advanced time; m is the externally measured mass), the blueshifted radiation density and the space-time curvature $\sim m_{\text{core}}/r_{\text{core}}^3$ near the core will grow to planckian values. At this stage one expects further growth to be damped by quantum processes. Now, for a black hole formed in a generic collapse, the inner-horizon radius r_{core} is not much smaller than the Schwarzschild radius $2m$. Thus, the inflow of blueshifted radiation has inflated the effective gravitational mass of the core to $m_{\text{core}} \approx m^3$ in Planck units. (Radiative scattering off the large curvature will disperse this mass more or less uniformly through the core.) Inside a black hole of five solar masses ($m = 10^{39}$) the core mass will inflate to 10^{57} times the mass of the observable universe.

We stress two points. First, no trace of this core inflation is perceptible to external observers, who continue to register the conventional mass m . This is essentially because news of the drastic internal change propagates as a gravitational wave, travelling at the speed of light, and thus can never escape from the hole. Second, core inflation requires a catalyst—in this case provided chiefly by radiation flowing out of the collapsing star. An object thrown down a deep well increases its material (kinetic) energy, but not its effective gravitating mass; the energy increase is balanced by a corresponding loss of gravitational potential energy, which itself acts as a gravitational source. The increased energy can manifest itself gravitationally only when the object is released from its gravitational binding. For a black hole, this is done by removing the well from the vicinity of the object! Under transverse irradiation from the star, the inner apparent horizon (the analogue of the well) is forced to contract and separate from the Cauchy horizon, allowing gravitational effects of the blueshifted inward stream of energy to become manifest.

The angular momentum of the core does not change as drastically (in an axisymmetric process it is conserved), and it becomes

* Present address: Physics Department, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1.

insignificant in comparison with the inflated mass. It seems plausible¹⁸ that non-spherical perturbations similarly become negligible on small angular scales. Thus, the geometry near the core closely resembles a Schwarzschild space-time of very large mass

$$ds^2 = -f^{-1}(r) dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) + f(r) dt^2 \quad (1)$$

where $f(r) = 2m_{\text{core}}/r - 1$ and $r \rightarrow r_{\text{core}} \approx m_{\text{core}}^{1/3}$.

Because f is positive, r here plays the part of a time, and t is a space-like (cylindrical) coordinate. In the case of a hole formed by a collapse, each instantaneous three-space—in particular, the core boundary where densities become planckian—may be visualized as a ‘fat cigar’, a cylinder extending infinitely along the positive t axis, and closed off at the end that enters the collapsing star by a cap that tapers to a point at the centre.

According to equation (1), a classical estimate of the proper time required to drop from the core boundary to the crunch at $r = 0$ is

$$\tau = \int_0^{r_{\text{core}}} f^{-1/2}(r) dr \approx 1$$

in Planck units. (This estimate is insensitive to the exact distribution of mass in the core.) The effect of blueshifted mass accretion by the core is therefore to foreshorten the proper time to the crunch from $r_{\text{core}} \approx m$ (about a day for a black hole of 10^{10} solar masses) to 10^{-43} s. This result is crucial for what follows.

The internal peculiarities of black holes have dramatic effects on the last minutes of a collapsing universe. A year before the crunch, galaxies begin to overlap, and the cosmic background radiation becomes hotter than the inside of a star: the stars break up and dissolve into the interstellar medium^{19,20}. By an hour or less to crunch time, supermassive black holes in galactic nuclei begin to merge^{19,20}. Within minutes, as the cores come together, their enormous gravitational masses become externally effective for the first time. The universe undergoes a phase transition (without substantial change of volume) to a state of planckian density and curvature. Proper time to the crunch is suddenly shortened to $\sim 10^{-43}$ s.

The black holes now merging would have formed at various, much earlier times, but the time deflation brings all cores to within Planck units of a common zero of time.

The first integral $f(r) dt/d\tau = \text{constant}$ for geodesic motion (with proper time τ) in the geometry described by equation (1) shows that transverse velocities $g_{tt}^{1/2} dt/d\tau \approx m^{-1}$ along the core axes are very small. Relative velocities will therefore be correspondingly small during core mergers; yet merging is complete within a few Planck times. (Merging is less a ‘collision’ than a simultaneous overlay—akin to a double exposure—of all parts of the cores in the region where they overlap, an effect of time deflation.)

The universe at this time is filled with a network of intersecting cylindrical cores, with axes randomly oriented. Their merging makes the three-geometry isotropic. Because of the small relative velocities, the smoothness of individual cores on scales approaching $r_{\text{core}} \approx m$ will be retained to some degree even after merging. The inflated mass is predominantly in the form of blueshifted gravitons of very high frequency. Thus, a few Planck times before the crunch, the universe is close to a radiation-dominated Friedmann geometry and possesses a uniformity (of non-thermal origin) with a coherence length enormously larger than the horizon scale.

Consider a simplified model of a closed universe (or a comoving volume in an open universe) in the last stages of contraction, which contains radiation and N black holes, each of the same external mass m , core radius $r_{\text{core}} = m$, and inflated core mass $m_{\text{core}} = m^3$. Before merging, the volume and radiation density are given by $R^3 = R_0^3(-t)^{3/2}$ and $\rho_{\text{rad}} = t^{-2}$ (R_0 is a constant). Just before merging, $R^3 = Nr_{\text{core}}^3$, and the mass stored in radiation and black holes is roughly comparable: $\rho_{\text{rad}} R^3 \approx Nm$. It

follows that the time of merging is $t = -t_M \sim -m$ to within an order of magnitude.

At this time, the average entropy density $S_H = Nm^2/R^3 = m^{-1}$ stored in black holes is much larger than the radiative entropy density $S_{\text{rad}} = T^3 = m^{-3/2}$, although still far below Planck levels. Immediately after merging, however, the radiative entropy density rises to $S_{\text{rad}} \approx 1$. At the same time, the Bekenstein–Hawking formula for black-hole entropy (which, strictly, requires an asymptotically flat exterior) becomes meaningless. Thus the paradox encountered previously no longer arises: the total entropy density always remains sub-planckian. Deflation of the conventional pre-crunch time t_M means that the universe is catapulted straight into the Planck regime, with no time remaining for adiabatic compression to raise its anomalously large (suprathemal) entropy density $S(-t_M) = m^{-1}$ above Planck levels.

Immediately after merging, and with the pre-crunch time recalibrated to $t \approx -t_0$ ($t_0 \ll t_M$), the universe is poised on the threshold of the quantum-gravity epoch. (We anticipate that $t_0 \approx 1$, but for simplicity we pretend that t_0 is one or two orders of magnitude larger, that is, that merging is complete well before quantum-gravitational effects become dominant.) At this stage, the universe can still be approximately described as a radiation-dominated Friedmann model, with density $\rho \approx t_0^{-2}$, volume Nr_{core}^3 and mass inflated to $M \approx Nm^3 t_0^{-2}$.

For the subsequent quantum phase no reliable theory currently exists, and we enter the realm of speculation. We proceed by assuming the existence of a consistent quantum theory of gravity that permits non-singular passage of a sufficiently homogeneous contracting structure through an epoch of maximal compression. We further assume that the dynamics is governed by a Wheeler–DeWitt equation in which the effective hamiltonian for matter and gravitation is invariant under ‘time’-reversal. It is then not difficult to prove formally that transition amplitudes are symmetric in initial and final states²¹. The probability will be maximized for a transition from a coherent state (in the sense of a minimal uncertainty wavepacket) at $t = -t_0$ to a coherent state at $t = +t_0$ peaked about the same three-metric and with equal and opposite extrinsic curvature.

Under these assumptions, a homogeneous model will probably bounce elastically between $t = -t_0$ and $t = +t_0$. But this universe differs grossly from its condition just before its black holes merged at $t = -t_M$. At $t = +t_M$, for comparison, its mass and volume have inflated by a factor $m^{3/2}$. The entropy density $S(t_M) \approx m^{-3/2}$ (now dominated by radiation) is sharply reduced from $S(-t_M) \approx m^{-1}$ (dominated by black holes). Thus one can speak of ‘regeneration of order near the crunch’, although the Second Law is never violated: the total entropy always increases²².

The universe re-emerging from the bounce may be homogeneous on scales much larger than the horizon. Although this structure is much larger than the critical Jeans length for gravitational instability, its overall uniformity is protected from immediate break-up by thermal and quantum fluctuations, because their gravitational influence, propagating causally, is limited to sub-horizon scales. A pre-established large-scale uniformity of this kind points towards a possible resolution of the well-known ‘horizon problem’^{1,9}.

It seems conceivable that, if our Universe is closed, its origin can be traced back through a finite number of cycles² of descending mass to a first ‘baby’ universe large enough to spawn mini black holes whose evaporation time enabled them to survive up to the crunch. And perhaps here, in whatever preceded this first cycle, one could seek answers to such riddles as the genesis of the dimensionality and metric signature of space-time and the vanishing of the cosmological constant. We shall not venture into these issues here. It is not our purpose to espouse any cosmogonical hypothesis; from our standpoint, no such hypothesis is verifiable^{23,24}. We intend only to point out some interesting features of the cosmic-bounce mechanism. This

remains, in our view, a viable alternative to currently favoured theories. $\square$

Received 10 August; accepted 6 November 1990.

1. Leslie, J. *Universes* (Routledge, London, 1989).
2. Dicke, R. H. & Peebles, P. J. E. in *General Relativity* (eds Hawking, S. W. & Israel, W.) 504–517 (Cambridge University Press, 1979).
3. Barrow, J. D. & Tipler, F. J. *The Anthropic Cosmological Principle*, 248–249, 619–621 (Clarendon Press, Oxford, 1986).
4. Zel'dovich, Ya. B. & Novikov, I. D. *Relativistic Astrophysics* Vol. 2, 658–663 (University of Chicago Press, 1983).
5. Parker, L. & Wang, Y. *Phys. Rev. D* **42**, 1877–1883 (1990).
6. Penrose, R. *Ann. N.Y. Acad. Sci.* **571**, 249–264 (1989).
7. Penrose, R. *The Emperor's New Mind* 339 (Clarendon, Oxford, 1989).
8. Tolman, R. C. *Relativity, Thermodynamics and Cosmology* 435 (Clarendon, Oxford, 1934).
9. Silk, J. *The Big Bang* 311 (W. H. Freeman, San Francisco, 1980).
10. Dicus, D. A., Letaw, J. R., Teplitz, D. C. & Teplitz, V. L. *Sci. Am.* **248**(3), 90–101 (1983).
11. Linde, A. D. *Rep. Prog. Phys.* **47**, 925–986 (1986).
12. Bekenstein, J. D. *Int. J. theor. Phys.* **28**, 967–981 (1989).
13. Poisson, E. & Israel, W. *Phys. Rev. Lett.* **63**, 1663–1666 (1989).
14. Poisson, E. & Israel, W. *Phys. Rev. D* **41**, 1796–1809 (1990).
15. Kaufman, W. J. III *Black Holes and Warped Spacetime* 110 (W. H. Freeman, San Francisco, 1979).
16. Israel, W. *Sci. Prog. (Oxford)* **68**, 333–363 (1983).
17. Price, R. H. *Phys. Rev. D* **5**, 2419–2438 (1972).
18. Barrabès, C., Israel, W. & Poisson, E. *Class. Quantum Gravity* (in the press).
19. Rees, M. J. *Observatory* **69**, 193–198 (1969).
20. Dicus, D. A., Letaw, J. R., Teplitz, D. C. & Teplitz, V. L. *Astrophys. J.* **252**, 1–9 (1982).
21. Brown, J. D. & Martinez, E. A. *Phys. Rev. D* **42**, 1931–1943 (1990).
22. Price, H. *Nature* **340**, 181 (1989).
23. Ellis, G. F. R., preprint 1990/1, Department of Applied Mathematics (University of Cape Town, 1990).
24. Maddox, J. *Nature* **340**, 425 (1989).

ACKNOWLEDGEMENTS. We thank R. Balbinot, D. Pavon and E. Poisson for discussions. This work was supported by the Canadian Institute for Advanced Research and the Natural Sciences and Engineering Research Council of Canada.

Old pulsars in the low-density globular clusters M13 and M53

S. R. Kulkarni*, S. B. Anderson*, T. A. Prince* & A. Wolszczan†

* Division of Physics, Mathematics and Astronomy,
California Institute of Technology, Pasadena, California 91125, USA
† Arecibo Observatory, PO Box 995, Arecibo, Puerto Rico 00612, USA

MILLISECOND pulsars are conventionally assumed to be spun up through the action of binary companions, although some subsequently lose their companions and appear as isolated pulsars. Such objects should therefore be more numerous in dense stellar systems. We report here the surprising discovery of two pulsars in low-density globular clusters: one is a single 10-ms pulsar (1639+36) in M13 (NGC6205), the other a 33-ms pulsar (1310+18) in a 256-day binary in M53 (NGC5025). Their ages, inferred from their luminosities and constraints on their period derivatives, seem to be $\sim 10^9$ years, significantly greater than previously reported ages ($\leq 10^8$ years) of cluster pulsars¹. The implied birth rate is inconsistent with the conventional two-body tidal capture model^{2,3}, suggesting that an alternative mechanism such as tidal capture between primordial binaries and a reservoir of (hundreds of) primordial neutron stars may dominate the production of tidal binaries in such clusters^{1,4}. The period derivative of PSR1639+36 is surprisingly small, and may be corrupted by acceleration due to the mean gravitational potential of the cluster⁵.

We discovered the pulsars during a survey of globular clusters with the Arecibo 305-m radio telescope^{6,7}. The observations were conducted at a centre frequency of 430 MHz and bandwidth of 10 MHz using the observatory's digital correlator in a manner identical to that described in ref. 8. The data analysis was carried out on Caltech's 512-node NCUBE supercomputer.

The pulse profiles of the two pulsars are shown in Fig. 1. By comparison with PSR2127+11A (ref. 8), we estimate S_{430} , the 430-MHz flux density, of PSR1310+18 and PSR1639+36 to be

1 mJy and 3 mJy, respectively. (Owing to the declination-dependent sensitivity of the Arecibo telescope, the signal-to-noise-ratio of PSR1639+36 is no better than that of PSR1310+18). The corresponding 430-MHz radio luminosities, $L \equiv S_{430}d^2$, are ~ 340 mJy kpc² and 150 mJy kpc² (d is the distance (kpc) to the cluster).

We obtained timing data using the same hardware as that used in the discovery. In the case of PSR1639+36, the fast sampled data, recorded on magnetic tape, were synchronously folded using a workstation computer and the resulting profiles cross-correlated with a high-quality template to yield times of arrival (see ref. 9 for details). The software package TEMPO was used to transform the topocentric times of arrival to the barycentre from which the usual pulsar parameters (see Table 1) were obtained. For PSR1310+18, the pulsar and orbital parameters are based on a least-squares fit to the apparent barycentric periods (Fig. 2), assuming that the pulsar is at the centre of the cluster.

PSR1639+36 is located 6.8 arcsec west and 21.4 arcsec south of the optical centroid, well within the 45-arcsec core of the cluster (Table 1). In the absence of a timing position, all we can say is that PSR1310+18 is within 6 arcmin (the primary beam half-width at half maximum of the Arecibo telescope at 430 MHz) of the centre of the cluster. Given the rarity of pulsars with short periods, the spatial coincidences are strong circumstantial evidence that the two pulsars are associated with their respective clusters.

Although large numbers of pulsars are being discovered in globular clusters, the discovery of pulsars in M13 and especially M53 was unexpected because the usual models^{2,3} predict essentially no pulsars in such low-density clusters. This is not a statistical fluctuation because pulsars have been found in all but two clusters (M3 and Pal 2) of the six low-density clusters searched from the Arecibo telescope. Indeed, in one of them (M5) two pulsars have been reported¹⁰.

In the standard model, primordial neutron stars (those born in the earliest epoch of the cluster) tidally capture other cluster stars (of mass M and radius R and moving at a relative velocity v) at the volumetric rate¹¹

$$\Gamma = 3 \times 10^{-10} \kappa(M) \frac{R}{R_\odot} \frac{M}{M_\odot} \frac{n_{\text{ns}}}{10^2 \text{ pc}^{-3}} \times \frac{n}{10^4 \text{ pc}^{-3}} \frac{10 \text{ km s}^{-1}}{v} \text{ yr}^{-1} \text{ pc}^{-3} \quad (1)$$

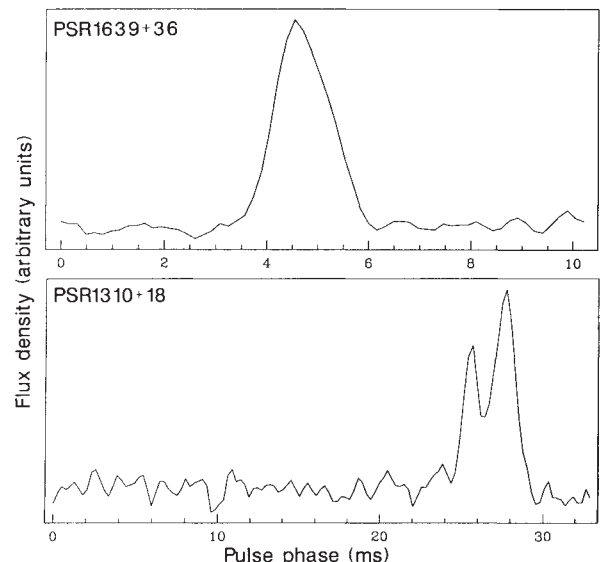


FIG. 1 The pulse profile of PSR1639+36 in M13 (top) and PSR1310+18 (bottom) at 430 MHz. The temporal resolution is limited by the sampling interval of 506 μ s.

Here n and n_{ns} are the volume densities of the non-degenerate cluster stars and the neutron stars and κ is essentially unity for $M \leq 0.8 M_{\odot}$. Mass transfer then spins up the neutron star and the end product is a spun-up pulsar in orbit around a white dwarf. In a competing model^{3,12}, a massive white dwarf is substituted for a neutron star. Copious mass transfer leads to the collapse of the accreting white dwarf into a neutron star after which the evolutionary path is the same as above. Main-sequence captures lead to binaries with orbital periods of about 1 day. About 15% of the encounters are with giants² of which only a quarter are expected to result in large-orbital-period (≥ 100 day) binaries, with the rest forming ultracompact binaries^{13,14}.

PSR1310+18 with an orbital period of 256 day and a mass function of $0.0098 M_{\odot}$ appears to be an approximate twin of PSR1620-26, the 11-ms 191-day binary pulsar in the cluster M4 (ref. 15). Both systems may be the result of a tidal capture of a giant star (but see discussion below). Assuming a mass of $1.4 M_{\odot}$ for PSR1310+18, the minimum mass of the companion is $0.3 M_{\odot}$ which is quite consistent with the expectation that it is a white dwarf.

The formation of a single pulsar such as 1639+36 is something of a puzzle. The simplest version of the standard model requires that the white dwarf/pulsar binary be disrupted by passing stars. In M28, a cluster that is denser than M13, this mechanism has already been found to be quite improbable¹⁶. We must therefore consider alternative mechanisms. The preponderance of single pulsars in clusters has already been noted^{1,17} and has been attributed¹ to the expansion of the companion on capture¹⁸, followed by the neutron star spiralling in and destroying the companion.

Integrating equation (1) over the volume of the cluster (assumed to be described by a King model with parameters from ref. 19 and assuming $M \approx 0.6 M_{\odot}$), we have estimated N_t , the number of two-body tidal captures over a Hubble lifetime ($t_H \approx 10^{10}$ yr) to be $4(N_{\text{ns}}/100)$ and $2(N_{\text{ns}}/100)$ in M13 and M53, respectively. Here, N_{ns} is the number of primordial neutron stars in the cluster. To get one wide binary pulsar, we need $N_{\text{ns}} = 1,300$. Even so, the expected age of such a pulsar in M53 would be $t_H/2$ (assuming constant formation rate over t_H) whereas below we argue for a smaller age, $\sim t_H/10$.

The birth rate of pulsars in these low-density systems seems to be quite high. In the absence of reliable values of the period derivative $\dot{P}$, we need other age estimators. Empirically it appears

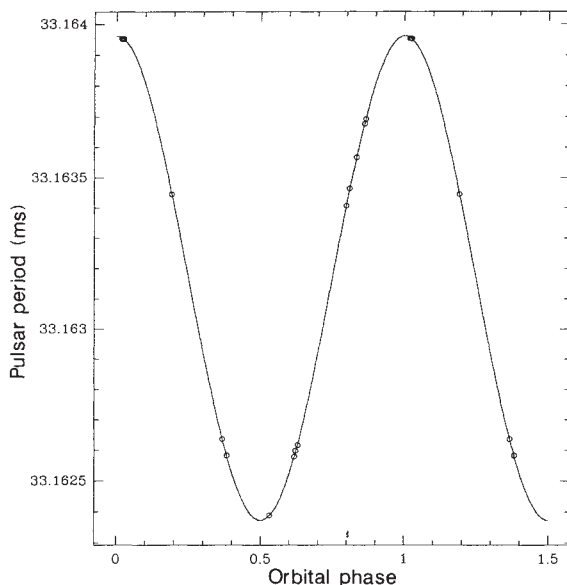


FIG. 2 The apparent period of PSR1310+18 as a function of the orbital phase.

TABLE 1 Pulsar and cluster parameters

Timing parameters: PSR1639+36*	
RA (B1950)	16 ^h 39 ^m 53.626(3) ^s
Dec (B1950)	36° 32' 54.94(6)''
Dispersion measure	30.36(4) cm ⁻³ pc
Barycentric period	10.3775094520(6) ms
Period derivative	<4.5 × 10 ⁻²⁰ s s ⁻¹ (5σ)
Epoch	2447666.71 JD
Globular cluster parameters: M13	
RA (B1950) ²³	16 ^h 39 ^m 54.19(5) ^s
Dec (B1950) ²³	36° 33' 16.3(6)''
Galactic latitude	59.0°
Galactic longitude	40.9°
Distance ¹⁹	7.1 kpc
Core Radius ¹⁹	45''
Timing parameters: PSR1310+18†	
Dispersion measure	24.0 ± 1.5 pc cm ⁻³
Barycentric period	33.163166(3) ms
Orbital period	255.84 ± 0.6 d
$a_1 \sin i$	84.173 ± 0.65 s
T_0	2447061.2 JD
e	<0.01
Globular cluster parameters: M53	
RA (B1950) ²³	13 ^h 10 ^m 28.27(5) ^s
Dec (B1950) ²³	18° 26' 02.3(6)''
Galactic latitude	332.9°
Galactic longitude	79.7°
Distance ¹⁹	18.5 kpc
Core Radius ¹⁹	22''

Numbers enclosed in parenthesis are the uncertainties in the last significant digits. Timing parameters derived from observed period data. The pulsar is assumed to be located at the centre of cluster. T_0 is the time of ascending node passage. RA, right ascension; dec, declination; JD, Julian day; $a_1 \sin i$, projected semi-major axis; e , eccentricity.

* Based on 17 observations made between 20 May 1989 and 19 May 1990. The r.m.s. of the post-fit residuals is 52 μs for an individual integration time of 23 min.

† Based on 13 observations made between 20 May 1989 and 8 April 1990 and one observation point on 29 Dec 1987. The r.m.s. after fitting to the velocity curve is 18 m s⁻¹.

that the radio luminosity of the disk millisecond pulsars decreases with the characteristic age (E. S. Phinney, S. R. K. and H. M. Johnston, manuscript in preparation) and using this relation, the luminosities of these two pulsars suggest ages of $\sim 10^9$ yr. This in turn increases the birth rate by a factor of about five and can be accommodated only by $N_{\text{ns}} \geq 10^3$, as above. But this would result in the core being dominated by degenerate stars, a result not consistent with dynamical studies²⁰. Thus our discovery offers strong support to models^{1,4} that invoke other mechanisms such as three-body tidal collisions involving primordial binaries and an N_{ns} of ~ 400 . We also note that the giant binary in M53 has a natural explanation in this framework.

The observed period derivative of PSR1639+36 is $<4.5 \times 10^{-20} \text{ s s}^{-1}$, a 5σ upper limit. Thus the nominal lower limit to the characteristic age of the pulsar is 3.8×10^9 yr. $\dot{P}$ of PSR1639+36 is, however, expected to be corrupted by the line-of-sight acceleration a induced by the mean cluster gravitational potential⁵

$$a(r_{\perp}) = \frac{GM(<r)}{r^2} \frac{\sqrt{r^2 - r_{\perp}^2}}{r} \quad (2)$$

Here r is the radial distance of the pulsar from the centre of the cluster, $r_{\perp}$ is the impact parameter and $M(<r)$ is the mass contained within radius r . This acceleration corrupts the intrinsic period derivative by an amount $\dot{P}_c = a(r_{\perp})P/c$.

For the observed $r_{\perp}$ we have calculated $a(r_{\perp})$ by assuming a King model to derive $M(<r)$, and by assuming⁵ that the neutron-star volume density is proportional to ρ^q where q (~ 2) is the

ratio of the mass of a neutron star to the mass of a typical star that contributes to the light distribution. The median and extreme accelerations are found to be $a = \pm 2.9 \times 10^{-7} \text{ cm s}^{-2}$ and $\pm 4.0 \times 10^{-7} \text{ cm s}^{-2}$ which will perturb the period derivative by $\pm 10 \times 10^{-20} \text{ s s}^{-1}$ and $\pm 14 \times 10^{-20} \text{ s s}^{-1}$; the sign of the acceleration depends on whether the pulsar is located ahead or behind the centre of the cluster.

From the above discussion, we can draw one firm conclusion: the intrinsic period derivative, $\dot{P}_p = \dot{P} - \dot{P}_c$ is $< 19 \times 10^{-20} \text{ s s}^{-1}$. Thus the pulsar is older than $0.9 \times 10^9 \text{ yr}$, about equal to the rough estimate of the age from our empirical luminosity

relation. This makes PSR1639+36 the oldest known cluster pulsar. A subsidiary conclusion is that the dipole field strength of the pulsar, estimated the usual way, $B_p = (B/10^9 \text{ G}) = \sqrt{P\dot{P}_{-15}}$, cannot be larger than 1.2 in which case the pulsar was spun up to no more than 4 ms, assuming that the equilibrium spin-up period is $3.45 B_p^{6/7} \text{ ms}$ (ref. 21).

The gradient of the mean gravitational field and the gravitational attraction due to the nearest neighbours will result in changes in the observed values of the higher period derivatives²². Measurement of $\dot{P}$, the largest effect, would unfortunately need six decades of timing observations. $\square$

Received 10 August; accepted 9 November 1990.

1. Phinney, E. S. & Kulkarni, S. R. *Nature* (submitted).
2. Verbunt, F. in *Neutron Stars & their Birth Events*, NATO ASI Ser. (ed. Kundt, W.) 179–218 (Kluwer, Dordrecht, 1990).
3. Bailyn, C. D. & Grindlay, J. E. *Astrophys. J.* **353**, 159–167 (1990).
4. Phinney, E. S. & Sigurdsson, S. *Nature* (in the press).
5. Phinney, E. S. *Mon. Not. R. astr. Soc.* (in the press).
6. Anderson, S., Kulkarni, S., Prince, T. & Wolszczan, A. *IAU Circ. No.* 4819 (1989).
7. Anderson, S., Kulkarni, S., Prince, T. & Wolszczan, A. *IAU Circ. No.* 4853 (1989).
8. Wolszczan, A. *et al. Nature* **337**, 531–533 (1989).
9. Anderson, S. B., Gorham, P. M., Kulkarni, S. R., Prince, T. A. & Wolszczan, A. *Nature* **346**, 42–44 (1990).
10. Wolszczan, A., Anderson, S., Kulkarni, S. & Prince, T. *IAU Cir. No.* 4880 (1989).
11. Lee, H. M. & Ostriker, J. P. *Astrophys. J.* **310**, 176–188 (1986).
12. Michel, C. F. *Nature* **329**, 310–312 (1987).
13. Bailyn, C. D. *Nature* **332**, 330–332 (1988).
14. McMillan, S. L. W., Taam, R. E. & McDermott, P. N. *Astrophys. J.* **354**, 190–200 (1990).

15. Lyne, A. G., Biggs, J. D., Brinklow, A., Ashworth, M. & McKenna, J. *Nature* **332**, 45–47 (1988).
16. Rappaport, S., Putney, A. & Verbunt, F. *Astrophys. J.* **345**, 210–221 (1989).
17. Kulkarni, S. R., Goss, W. M., Wolszczan, A. & Middleditch, J. M. *Astrophys. J.* **363**, L5–L8 (1990).
18. Ray, A., Kembhavi, A. K. & Antia, H. M. *Astr. Astrophys.* **184**, 164–172 (1987).
19. Chernoff, D. F. & Djorgovski, S. D. *Astrophys. J.* **339**, 904–918 (1989).
20. Meylan, G. *Astr. Astrophys.* **206**, 106–112 (1989).
21. Bhattacharya, D. & van den Heuvel, E. J. P. *Phys. Rep.* (in the press).
22. Blandford, R. D., Romani, R. W. & Applegate, J. H. *Mon. Not. R. astr. Soc.* **225**, 51P–53P (1987).
23. Shawl, S. J. & White, R. E. *Astr. J.* **91**, 312–316 (1986).

ACKNOWLEDGEMENTS. We thank A. S. Fruchter for donation of telescope time, J. H. Taylor for the TEMPO software package, and E. S. Phinney and S. Djorgovski for discussions. This work was supported by NSF Presidential Young Investigator awards (T.A.P. and S.R.K.), the Department of Energy (T.A.P.), an NSF Graduate Fellowship (SBA), and an Alfred P. Sloan Fellowship (S.R.K.). Arecibo Observatory is part of the National Astronomy and Ionosphere Centre, operated by Cornell University under contract with the NSF.

Seismological measurement of solar helium abundance

S. V. Vorontsov*, V. A. Baturin† & A. A. Pamyatnykh‡

* Institute of Physics of the Earth, Moscow 123810, USSR

† Sternberg Astronomical Institute, Moscow 119899, USSR

‡ Astronomical Council, Moscow 109017, USSR

THE internal structure and evolution of the Sun depends on its chemical composition, particularly the helium abundance. In addition, the helium abundance in the solar envelope is thought to represent the protosolar value, making it a datum of cosmological significance. Spectroscopic measurements of the helium abundance are uncertain, and the most reliable estimates until now have come from the calibration of solar evolutionary models. The frequencies of solar acoustic oscillations are sensitive, however, to the behaviour of the speed of sound in the Sun's helium ionization zone, which allows a helioseismological determination of the helium abundance. Sound-speed inversion of helioseismological data can be used for this purpose^{1–3}, but precise frequency measurements of high-degree oscillation modes are needed. Here we describe a new approach based on an analysis of the phase shift of acoustic waves^{4,5} of intermediate-degree modes. From the accurate intermediate-mode data now available⁶, we obtain a helium mass fraction $Y = 0.25 \pm 0.01$ in the solar convection zone, significantly smaller than the value $Y = 0.27–0.29$ predicted by recent solar evolutionary models^{7–9}. The discrepancy indicates either that initial helium abundance was reduced in the envelope by downward diffusion⁹ or that the protosolar value was lower than currently accepted.

The sensitivity of the low- and intermediate-degree solar oscillation frequencies to helium abundance is very small¹⁰: effects of the order of 10^{-4} , close to the accuracy of frequency measurements, are significant for our study. The frequencies themselves depend in a rather sophisticated way on the solar structure. A large number of oscillation frequencies have been measured, but the information contained in the frequencies of different modes is not independent. The main properties of these high-frequency acoustic modes are reproduced by the simple

asymptotic theory. (For reviews of solar seismology, see refs 4 and 11.)

Our technique consists of two steps. We first describe the observational data by an approximation using many fewer parameters than the number of frequencies, thus improving the signal-to-noise ratio. From this approximation we define some functions, or 'tracers', sensitive to helium abundance. We then calibrate the helium abundance by fitting these 'tracers' with those calculated for a variety of solar-envelope models.

We approximate the observational data by functions $F(w)$ and $\alpha(w)$ in the eigenfrequency equation that results from the asymptotic analysis⁴

$$\frac{\omega}{\pi} F(w) - \alpha(w) \approx n(w, \omega)$$

where ω is the angular frequency, $w = \omega/(l + \frac{1}{2})$ determines the penetration depth, l is the degree, and n is the radial order of

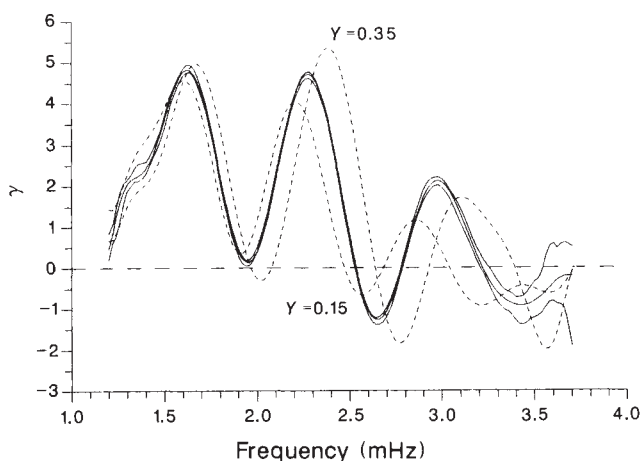
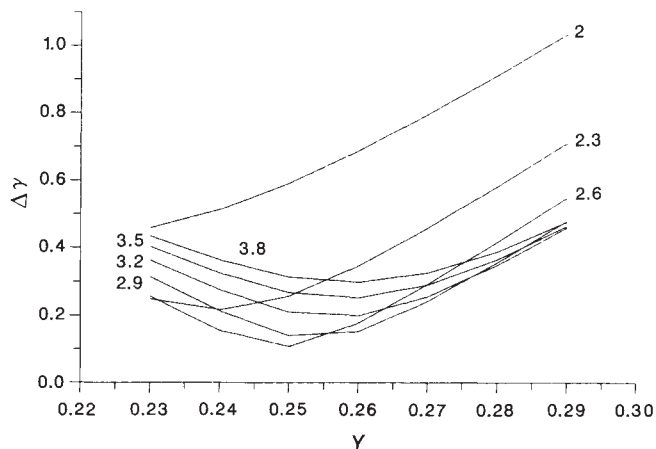


FIG. 1 Phase function $\gamma(\omega)$ inferred from the observational frequencies. The three solid lines show the best-fit approximation and the envelope of the solutions obtained by adding white noise to the observational frequencies, with limiting amplitude equal to the reported observational errors, in 100 runs, for polynomials of order 20, 21, 22. Dashed lines show $\gamma(\omega)$ calculated for two models with $Y=0.15$ and $Y=0.35$.

FIG. 2 Root-mean-square difference between $\gamma(\omega)$ calculated for envelope models with different helium abundance Y and convection efficiency, and $\gamma(\omega)$ inferred from the observational frequencies. Each line is labelled with the mixing-length parameter (ratio of mixing length to pressure scale-height), which determines the specific entropy in the helium ionization zone.



an acoustic mode. $F(\omega)$ is determined by the speed of sound in the solar interior and $\alpha(\omega)$ is the phase shift that arises when an internal acoustic wave is reflected by the surface layers. The right-hand side of equation (1) represents observational data by radial order n , as a function of two variables w and ω .

We used intermediate-degree modes trapped in the solar convection zone; these modes are accurately described by equation (1). Specifically, we selected modes in the w range limited by $l=50$ and $l=100$ at a frequency of 3 mHz (a total of 594 modes from individual frequency measurements⁶). We approximated $F(\omega)$ and $\alpha(\omega)$ by polynomials^{5,12}; polynomials of order 20 for $\alpha(\omega)$ and order 7 for $F(\omega)$ were found appropriate. The root-mean-square residual of the approximation of $n(w, \omega)$ is 1×10^{-4} .

For calibration of the envelope model, we define 'tracers':

$$\gamma(\omega) = \alpha(\omega) - \omega \frac{d\alpha}{d\omega} - \omega^2 \frac{d^2\alpha}{d\omega^2}$$

$$\beta(\omega) = \alpha(\omega) - \omega \frac{d\alpha}{d\omega}$$

$$n(\omega) = n(w_0, \omega) = \frac{\omega}{\pi} F(w_0) - \alpha(\omega)$$

where w_0 is some mean value of w . These functions can be calculated for a given envelope model without calculating the oscillation frequencies^{4,5,13}. The functions are determined by the structure of the envelope alone, so that it is not necessary to calculate the full evolution of solar models. The calculations are thus very fast, and large numbers of envelope models can be tested.

The most sensitive 'tracer' of helium abundance is $\gamma(\omega)$. In the theoretical description of solar p -modes (in which pressure is the restoring force) this function accounts for the deviations of the exact solutions of wave equations from the second-order asymptotic approximation¹³. The second helium ionization zone, with a thickness of the order of the radial wavelength⁴, produces strong quasiperiodic fluctuations in $\gamma(\omega)$ and the amplitude and phase of this periodicity are sensitive to helium abundance (Fig. 1).

The results presented in Figs 2 and 3 were obtained with envelope models calculated using the atmospheric temperature profile of the Harvard-Smithsonian reference atmosphere¹⁴ and a standard formulation of the mixing-length theory¹⁵. We used the opacity tables of Cox and Tabor¹⁶, but in the outer layers the opacities were increased roughly in accordance with refs 17 and 18. The equation of state was calculated by solving the Saha ionization equations, with electrostatic corrections included in the Debye-Hückel approximation^{19,20}. We assumed a heavy-element abundance of $Z = 0.02$.

The models were calibrated by fitting the theoretical and observational functions $\gamma(\omega)$, $\beta(\omega)$ and $n(\omega)$ in the frequency

range 1.5–2.5 mHz. At these low frequencies, the influence of atmospheric uncertainties and possible non-adiabatic effects⁹ are relatively small. Figure 2 shows the calibration of $\gamma(\omega)$. Calibrations of $\beta(\omega)$ (similar to those used in refs 12 and 21) and $n(\omega)$ (which is essentially the calibration for the frequencies themselves) are less sensitive to helium abundance, but the best-fit models are roughly the same. Figure 3 shows the result of the calibration using all three functions together. We estimate the helium abundance to be $Y = 0.25 \pm 0.01$.

To test the accuracy of the theoretical description, we calculated the oscillation frequencies for the envelope models, and determined the functions used in the calibration from the frequencies in the same way as for the observational frequencies. The helium abundance in the best-fit models is shifted to slightly lower values, within the limits given above. We also tested similar models, but with different atmospheres²² and with lower opacities (taken from ref. 16 without modifications); the helium abundance was unchanged.

The main problem with the calibration is connected with physical uncertainties in the equation of state. We tested four versions: (1) A simple Saha equation of state. We could not bring the models into reasonable agreement with observational data (see also ref. 12), so this equation of state is oversimplified. (2) A Saha equation with electrostatic corrections—the results were discussed above. (3) An equation of state with the partition functions calculated in accordance with a version of the 'confined atom' model of Däppen, which includes electrostatic corrections²³. (We scaled the number of bound states as a function of mean volume per particle so that the partition functions did not increase with depth.) The helium abundance

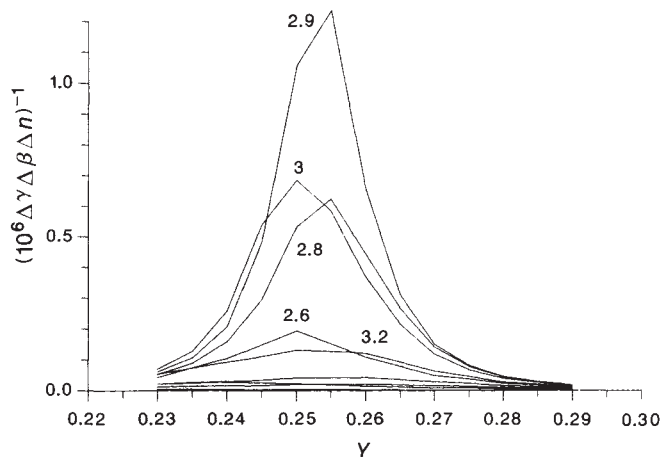


FIG. 3 The reciprocal of $\Delta\gamma\Delta\beta\Delta n$, where $\Delta\gamma$, $\Delta\beta$ and Δn are root-mean-square differences between $\gamma(\omega)$, $\beta(\omega)$ and $n(\omega)$ calculated for the envelope models and those inferred from the observational data. The maximum of the curves corresponds to the best-fit model.

remained the same. (4) The most recent, so-called MHD (Mihalas, Hammer, Däppen) equation of state^{19,20}. Although we only used three mixtures with $Y = 0.23, 0.26$ and 0.29 in the models (see ref. 12), the amplitude and phase behaviour of $\gamma(\omega)$ indicate that the helium abundance is within the same limits. The depth of the convection zone for the best-fit models is found to be in agreement with that determined by the sound-speed inversions ($30 \pm 1\%$ of the solar radius²⁴).

Our result for the helium abundance in the convection zone is significantly lower than the initial solar helium abundance, $Y = 0.27\text{--}0.29$, predicted by the recent evolutionary models^{7–9}. This discrepancy may be connected with some oversimplified assumptions in the standard evolutionary calibrations. Possible helium diffusion from the convection zone into the radiative interior during solar evolution⁹ may result in a lowering of the helium abundance in the envelope from its initial value. The discrepancy may also be associated with standard assumptions about the composition and evolution of the solar core. Indeed, recent helioseismic studies reveal some problems with the seismic structure of the solar core that is predicted by the evolutionary models^{24–26}. In the near future, more accurate seismic sound-ing of the deep solar interior should help clear up the uncertainties about the seismic structure of the core.

Two independent helioseismic measurements of helium abundance were performed recently using linearized inversion techniques. One of the results is somewhat higher, and the other is even lower than our value, inferred by direct calibration of the envelope model. Using the optimal averaging procedure of Backus and Gilbert, Gough *et al.*^{27,28} obtained $Y = 0.268 \pm 0.002$. Using a least-squares spline-fitting technique, Dziembowski *et al.*²⁹ found $Y = 0.235 \pm 0.005$. Although all these results give a lower helium abundance than the evolutionary models, the reasons for the discrepancies between them remain to be studied. *Note added in proof.* A similar result, $Y \approx 0.25$, was obtained recently by Christensen-Dalsgaard and Pérez Hernández³⁰ by analysing the acoustic phase shift using a different technique. $\square$

Received 14 August; accepted 9 November 1990.

- Gough, D. O. *Mem. Soc. astr. Ital.* **55**, 15–35 (1984).
- Däppen, W. & Gough, D. O. In *Theoretical Problems in Stellar Stability and Oscillations* (eds Gabriel, M. & Noels, A.) 264–269 (Institute d'Astrophysique, Liege, 1984).
- Däppen, W., Gough, D. O. & Thompson, M. J. in *Seismology of the Sun and Sun-Like Stars*, ESA SP-286 (ed. Rolfe, E. J.) 505–510 (European Space Agency, Noordwijk, 1988).
- Vorontsov, S. V. & Zharkov, V. N. *Sov. Sci. Rev. E7*, 1–103 (1989).
- Brody, M. A. & Vorontsov, S. V. in *Seismology of the Sun and Sun-Like Stars*, ESA SP-286 (ed. Rolfe, E. J.) 487–491 (European Space Agency, Noordwijk, 1988).
- Libbrecht, K. G., Woodard, M. F. & Kaufman, J. M. *Astrophys. J. Suppl.* (in the press).
- Lebreton, Y. & Maeder, A. *Astr. Astrophys.* **161**, 119–124 (1986).
- Turck-Chieze, S., Cahen, S., Casse, M. & Doom, C. *Astrophys. J.* **335**, 415–424 (1988).
- Cox, A. N., Guzik, J. A. & Kidman, R. B. *Astrophys. J.* **342**, 1187–1206 (1989).
- Guenther, D. B. & Sarajedini, A. *Astrophys. J.* **327**, 993–997 (1988).
- Christensen-Dalsgaard, J., Gough, D. & Toomre, J. *Science* **229**, 923–931 (1985).
- Pamyatnykh, A. A., Vorontsov, S. V. & Däppen, W. *Astr. Astrophys.* (in the press).
- Vorontsov, S. V. *Astr. Zh.* (in the press).
- Gingerich, O., Noyes, R. W., Kalkofen, W. & Cuny, Y. *Solar Phys.* **18**, 347–365 (1971).
- Paczynski, B. *Acta astr.* **19**, 1–22 (1969).
- Cox, A. N. & Tabor, J. E. *Astrophys. J. Suppl.* **31**, 271–312 (1976).
- Cox, A. N. in *Inside the Sun* (eds Berthomieu, G. & Cribier, M.) 61–80 (Kluwer, Dordrecht, 1990).
- Christensen-Dalsgaard, J. in *Inside the Sun* (eds Berthomieu, G. & Cribier, M.) 305–326 (Kluwer, Dordrecht, 1990).
- Mihalas, D., Däppen, W. & Hummer, D. G. *Astrophys. J.* **331**, 815–825 (1988).
- Däppen, W., Mihalas, D., Hummer, D. G. & Mihalas, B. W. *Astrophys. J.* **332**, 261–270 (1988).
- Baturin, V. A. *Pis'ma Astr. Zh.* (in the press).
- Vernazza, J. E., Avrett, E. H. & Loesser, R. *Astrophys. J. Suppl.* **45**, 635–725 (1981).
- Däppen, W. *Astr. Astrophys.* **91**, 212–220 (1980).
- Vorontsov, S. V. in *Seismology of the Sun and Sun-Like Stars*, ESA SP-286 (ed. Rolfe, E. J.) 475–480 (European Space Agency, Noordwijk, 1988).
- Gough, D. O. & Kosovichev, A. G. in *Inside the Sun* (eds Berthomieu, G. & Cribier, M.) 327–340 (Kluwer, Dordrecht, 1990).
- Dziembowski, W. A., Pamyatnykh, A. A. & Sienkiewicz, R. *Mon. Not. R. astr. Soc.* **244**, 542–550 (1990).
- Gough, D. O., Kosovichev, A. G. & Thompson, M. J. in *Challenges to Theories of the Structure of Moderate-Mass Stars* (eds Gough, D. O. & Toomre, J.) (in the press).
- Gough, D. & Toomre, J. in *GONG Newsletter No. 14* (National Solar Observatory, Tucson, 1990).
- Dziembowski, W. A., Pamyatnykh, A. A. & Sienkiewicz, R. *Mon. Not. R. astr. Soc.* (submitted).
- Christensen-Dalsgaard, J. & Pérez Hernández, F. in *Challenges to Theories of the Structure of Moderate-Mass Stars* (eds Gough, D. O. & Toomre, J.) (in the press).

ACKNOWLEDGEMENTS. We thank K. G. Libbrecht for providing observational data. Envelope models with the MHD equation of state were calculated in collaboration with W. Däppen. The work was partly supported by 'Zodiac' Scientific Methodological Council.

Large amounts of extinct ^{26}Al in interstellar grains from the Murchison meteorite

Ernst Zinner*, Sachiko Amari*†, Edward Anders† & Roy Lewis†

*McDonnell Center for the Space Sciences and the Physics Department, Washington University, St. Louis, Missouri 63130-4899, USA

† Enrico Fermi Institute and Department of Chemistry, University of Chicago, Chicago, Illinois 60637-1433, USA

INTERSTELLAR graphite and silicon carbide grains recovered from the Murchison CM2 chondritic meteorite are known to show large anomalies in the isotopic abundances of neon, xenon, carbon, nitrogen and silicon^{1–3}. These anomalies provide clues to the nucleosynthetic origin of the material from which the grains formed. Here we report that both types of grain also have large abundances of ^{26}Mg from the decay of extinct ^{26}Al (half-life 705,000 years). The deduced initial $^{26}\text{Al}/^{27}\text{Al}$ ratios range up to 0.06 in graphite and 0.2 in SiC—1,200 and 4,000 times the maximum values found in refractory inclusions in primitive meteorites. All proposed stellar sources of carbonaceous dust (red giants, novae, Wolf-Rayet stars and supernovae) also produce ^{26}Al , but the highest $^{26}\text{Al}/^{27}\text{Al}$ ratios found in these grains seem to rule out Wolf-Rayet stars and supernovae. The aluminium abundance correlates with that of nitrogen, suggesting that the aluminium condensed as aluminium nitride.

Extending our search for isotope anomalies to minor elements, we measured, using an ion microprobe, the Mg and Al in silicon carbide and 'carbon- α ' (spherulitic graphite) separates from the Murchison meteorite. Four SiC separates of the following nominal grain sizes were analysed: HM (0.03–0.2 μm), HN (0.2–2 μm)³, KJG (1.5–3.0 μm)⁴, and LS and LU (henceforth designated LS-U) (>2 μm)⁵. In the first two, we were able to measure only aggregates of many grains, but in the last two, we studied individual grains ranging from 1.6 to 6.0 μm for KJG and from 3 to 23 μm for LS-U. The morphology and major-element chemistry of all single grains was determined by scanning electron microscopy/energy dispersive X-ray (SEM-EDX) analysis, and their C, N and Si isotopic compositions were measured by ion-microprobe mass spectrometry. Complete data will be reported elsewhere; the compositions of a few unusual grains are given in Table 1. Individual Ca grains were from separate KFA1 (density 2.05–2.10 g cm^{-3} , $^{22}\text{Ne-E(L)} = 13,000 \times 10^{-8} \text{ cm}^3 \text{ g}^{-1}$)⁶. This separate consists of carbon grains >1 μm in size, the majority being spherical and either isotopically heavy or normal⁷. We selected it because the correlation of minor elements (H, N, Si) with heavy carbon⁸ offered a better chance to find Al in these grains. Still, we had to restrict ourselves to the largest grains (>2.5 μm).

Because ^{26}Al decays to ^{26}Mg , its signature is an increase in $^{26}\text{Mg}/^{24}\text{Mg}$ over the normal (Solar System) ratio. Thus $^{26}\text{Mg}_{\text{excess}}/^{27}\text{Al} = (^{26}\text{Al}/^{27}\text{Al})_0$ at $t = 0$, when the grain formed. In contrast to C, N and Si, Mg and Al were measured as positive secondary ions. In KJG, Si isotope ratios were also obtained from positive ions. We used terrestrial SiC, spinel and augite as isotope standards and to obtain relative sensitivity factors for the elements measured. The $^{26}\text{Al}/^{27}\text{Al}$ ratios and Al concentrations given in the table and figures are based on these factors. All measurements on HM (7 analyses) and HN (17 analyses) gave clear ^{26}Mg excesses, with $^{26}\text{Al}/^{27}\text{Al}$ ratios ranging from 7×10^{-4} to 2.2×10^{-3} in HM and from 1.2×10^{-3} to 3.2×10^{-2} in HN. In many cases these ratios are only lower limits because these separates contain Al-rich oxides such as corundum (Al_2O_3) and hibonite ($\text{Al}_{12}\text{CaO}_{19}$).

TABLE 1 Isotopic compositions of selected grains

Grain	Al (wt%)	$^{26}\text{Al}/^{27}\text{Al}$	$^{12}\text{C}/^{13}\text{C}$	$^{14}\text{N}/^{15}\text{N}$	$\delta^{29}\text{Si}$ (‰)	$\delta^{30}\text{Si}$ (‰)
1 SiC-LS	2.08	0.0021	3.00	101	-15.9	-9.7
2 SiC-LU	1.12	0.0039	4.59	1,690	101.7	63.8
3 SiC-LU	0.090	<0.00014	48.4	680	43.4	44.3
4 SiC-KJG	2.24	0.20	1,135	18.1	-379	-591
5 SiC-KJG	2.63	0.015	54.3	209	16.6	6.4
6 SiC-KJG	0.80	0.011	54.2	214	19.2	23.2
7 SiC-KJG	0.17	0.0016	14.0	747	33.7	45.0
8 C-KFA1	0.39	0.063	47.8	n.m.	n.m.	n.m.
9 C-KFA1	0.46	0.037	134	n.m.	n.m.	n.m.
10 C-KFA1	0.063	0.061	88.3	n.m.	n.m.	n.m.
11 C-KFA1	0.15	0.0065	8.9	n.m.	n.m.	n.m.
12 C-KFA1	0.17	0.014	47.5	n.m.	n.m.	n.m.

$\delta^{29}\text{Si} = 1,000 \times [(^{29}\text{Si}/^{28}\text{Si})_{\text{sample}} / (^{29}\text{Si}/^{28}\text{Si})_{\text{standard}} - 1]$. Standard: synthetic SiC.
n.m., not measured.

The two separates LS-U and KJG contain very different populations of SiC. The large grains of the first are anhedral with smooth surfaces. In contrast, the smaller grains of KJG have a platy morphology, often showing hexagonal features, although Raman spectroscopy indicates a cubic lattice structure (B. Wopenka, personal communication). The two groups also differ in their N and Al contents (Fig. 1) and their isotopic compositions. With three exceptions, LS-U grains have low N and Al, and barely overlap with the KJG distribution. The general correlation between N and Al suggests that Al in SiC is present as AlN (ref. 9), in solid solution (SiC and AlN are isoelectronic) or as AlN inclusions. In the latter case these inclusions must be small and fairly uniformly distributed in the SiC grains because we did not observe large fluctuations in CN^-/C^- and Al^+/Si^+ during analysis. Some of the scatter in Fig. 1 must be due to contamination, but some could be due to varying sizes of AlN inclusions because formation of CN^- requires nearby C atoms.

LS-U has a restricted range of C isotope compositions. With the exception of 3 grains out of 41 (1 and 2 in Table 1 and one with $^{12}\text{C}/^{13}\text{C} = 19.5$) $^{12}\text{C}/^{13}\text{C}$ ratios lie between 47.0 and 56.5, with a few grains being close to normal (89). KJG grains have a much wider distribution, between $^{12}\text{C}/^{13}\text{C} = 5.8$ and 155. One KJG grain (4 in Table 1) has the most extreme C, N and Si isotope ratios seen so far in SiC.

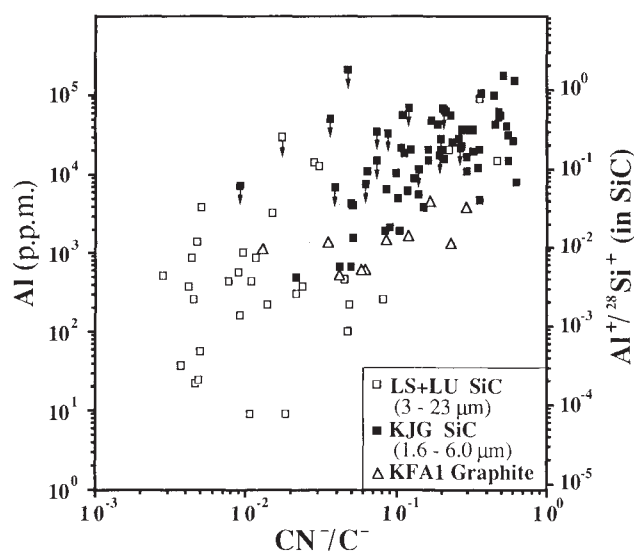


FIG. 1 Al content in interstellar SiC and graphite grains plotted against the CN^-/C^- ratio. The rough correlation indicates that Al exists as AlN in interstellar SiC and graphite grains. The large SiC grains ($>3 \mu\text{m}$) of LS-U are distinguished from the smaller (1.6–6.0 μm) grains of KJG by their smaller N and Al contents. Because of uncertainties in the relative sensitivity factors the Al concentration scale is only approximate.

The Mg and Al data from single grains are shown in Fig. 2, and again there is a marked difference between the two populations. Only 2 LS-U grains (out of 32) have clear excesses of ^{26}Mg , the others gave only upper limits on the $^{26}\text{Al}/^{27}\text{Al}$ ratio. The corresponding fraction is much higher for KJG (40 out of 70), owing mainly to the higher Al concentrations rather than higher intrinsic $^{26}\text{Al}/^{27}\text{Al}$ ratios in the smaller grains. With one exception ($\delta^{25}\text{Mg} = -250 \pm 80\%$) the $^{25}\text{Mg}/^{24}\text{Mg}$ ratios are normal within the sometimes fairly large errors ($^{24}\text{Mg}^+$ count rates were as low as 0.06 count s^{-1}) whereas $^{26}\text{Mg}/^{24}\text{Mg}$ vary. The highest $^{26}\text{Mg}/^{24}\text{Mg}$ ratios were 600 and 966 in KJG grains (up to $7,000 \times$ the Solar System ratio¹⁰) and 5.3 and 10 in KFA1 grains. This shows that the ^{26}Mg excesses come from the decay of ^{26}Al inside the SiC or graphite, and that Al was incorporated preferentially over Mg.

Such preferential trapping of Al is expected during condensation of SiC in stellar outflows. Thermodynamic calculations⁹ for a gas with $\text{C}/\text{O} > 1$, but otherwise solar composition, show that AlN condenses after SiC but well before Mg phases; thus volatility, as well as structural similarity, favours the uptake of AlN in solid solution. For solar abundances, complete condensation of Al would give 5.7×10^4 p.p.m. Al in SiC, as shown by the dashed line in Fig. 2.

There are some correlations between the ^{26}Al content and the $^{12}\text{C}/^{13}\text{C}$ ratio. The two LS-U grains with detectable ^{26}Al have exceptionally small $^{12}\text{C}/^{13}\text{C}$ ratios and grain 2 also has an exceptionally large $^{14}\text{N}/^{15}\text{N}$ ratio. In Fig. 2 these two grains plot next to a cluster containing five out of the six KJG grains with the heaviest C ($^{12}\text{C}/^{13}\text{C}$ ratios between 5.8 and 14.0; all others have $^{12}\text{C}/^{13}\text{C} > 26.7$). The grains in this cluster have cosmic Al/Si. The three KJG grains with the largest $^{26}\text{Al}/^{27}\text{Al}$ have ^{15}N excesses (grains 4, 5 and 6) but 5 and 6 have fairly commonplace $^{12}\text{C}/^{13}\text{C}$ ratios of 54 (Table 1). These two grains have practically identical C, N and Si isotopes ($\delta^{30}\text{Si}$ differs by 1.9σ) and are either fragments of the same parent grain or formed at the same place. For the other KJG grains with detectable ^{26}Al there are no obvious relationships between $^{26}\text{Al}/^{27}\text{Al}$ and other isotope signatures.

Figures 1 and 2 also show the data on the ten graphite grains that we measured. Five contain detectable ^{26}Al , with $^{26}\text{Al}/^{27}\text{Al}$ ratios between 6.5×10^{-3} and 6.3×10^{-2} . The Al contents are lower than in the SiC grains with the highest $^{26}\text{Al}/^{27}\text{Al}$, but approach the cosmic Al/C ratio of 8×10^3 p.p.m. (for $\text{C}/\text{O} = 1$). The Al content is usually higher than those of Mg and Si despite its tenfold lower cosmic abundance, indicating that it was incorporated during condensation and not implanted. On average, $^{26}\text{Al}/^{27}\text{Al}$ in the carbon grains is higher than in SiC. This could, in part, be a selection effect because we concentrated on the largest grains. ^{26}Al does not correlate with $^{12}\text{C}/^{13}\text{C}$, being present in carbon grains with light, normal and heavy carbon (Table 1).

The association of Al with N in graphite (Fig. 1) has interesting

implications for the origin of circumstellar graphite. It is very likely that Al again condensed as AlN, but because AlN does not form solid solutions with graphite, it must be present as discrete grains. These grains must be well inside the graphite spherules, or they would not have survived the extensive chemical processing of our samples (AlN hydrolyses in water). Thus AlN grains must have been present while the graphite spherules were still growing, although the equilibrium condensation temperature of AlN, 1,040 K at 2×10^{-8} atm and C/O = 1.5 (ref. 9), is much lower than that of graphite (1,700 K). Perhaps graphite nucleation was kinetically inhibited^{11,12} and was triggered by the appearance of AlN grains. A key role for Al is suggested by its high abundance relative to other more abundant but less refractory metals: Al/Mg and Al/Si are up to 1,500 and 140 times the cosmic ratio. Heterogeneous nucleation of amorphous C on SiC has been proposed¹³, but our data suggest that AlN was responsible, although SiC is cosmically more abundant and condenses before AlN.

The most striking result is that $^{26}\text{Al}/^{27}\text{Al}$ in SiC and carbon grains far exceeds the canonical maximum value of 5×10^{-5} in Al-rich minerals from refractory inclusions of primitive meteorites¹⁴. There is, however, an important difference between these ratios. The Al-rich phases in the inclusions were formed from material that melted or was vaporized during formation of the Solar System, thereby losing radiogenic ^{26}Mg accumulated up to that time, and so their $^{26}\text{Al}/^{27}\text{Al}$ ratio reflects the amount of ^{26}Al then remaining. The SiC and C grains, on the other hand, apparently were not heated sufficiently to expel radiogenic ^{26}Mg , as both phases are destroyed in a hot solar gas and at least some of the grains still contain exotic noble gases⁴. Thus their $^{26}\text{Al}/^{27}\text{Al}$ ratio refers to their time of formation, not to their time of arrival in the Solar System. The possibility of such fossil ^{26}Mg in presolar grains has been emphasized by Clayton¹⁵.

The interval between these two events is given by the presolar cosmic-ray exposure age of the grains, as calculated from their cosmogenic ^{21}Ne content⁴. For KJG, the nominal age is 106 Myr, far too long to permit survival of ^{26}Al . This age is an average for old and young grains, however, and some of the young grains may still have contained ^{26}Al on arrival.

The corresponding age for graphite is not known, because carbon is too light a target element to produce ^{21}Ne . Again, because the graphite grains presumably formed in different stars at different times, some may have carried ^{26}Al into the Solar System. The amount of ^{26}Al contributed by SiC and graphite may have been much larger than implied by their present-day abundance in meteorites. Carbon stars are thought to produce two-thirds of all interstellar dust¹⁶, but much of this carbonaceous dust would be destroyed on entering the inner Solar System, because SiC and C are unstable in a hot gas with C/O < 1.

Not having been reset during formation of the Solar System, the $^{26}\text{Al}/^{27}\text{Al}$ ratios of SiC and graphite accurately reflect the production ratios in stars and thus tell us something about the stellar sources of ^{26}Al . For all stars considered, transport times from the production site to the surface are short compared to the half-life of ^{26}Al , and the timescale for ejection and grain formation is shorter still. Thus, there should be little decay between formation and condensation of ^{26}Al . But dilution by unprocessed material from the stellar envelope will vary, and therefore the highest ratios are the most significant.

Among the proposed production mechanisms for ^{26}Al are hydrostatic hydrogen burning¹⁷ in AGB red giants (of the asymptotic giant branch, AGB)¹⁸ or in massive stars during the Wolf-Rayet stage^{19,20}, explosive hydrogen burning in novae²¹⁻²³, and explosive carbon and neon burning in type II supernovae²⁴⁻²⁶. In principle, all of these stellar environments can have C > O, enabling the formation of graphite and SiC. Although dust formation in supernovae is still controversial, the other stellar sites are considered to be sources of carbon-rich dust grains²⁷. The $^{26}\text{Al}/^{27}\text{Al}$ production ratios predicted for the different mechanisms vary over several orders of magnitude, from $\sim 10^{-3}$ in supernovae to at least 1 in AGB stars and novae. Although the lower ratios in the SiC grains are compatible with supernovae and Wolf-Rayet stars, the highest ratios in SiC and graphite grains seem to rule out these two sources for those grains. We emphasize, however, that there are still large uncertainties in model calculations of ^{26}Al production²². At present, it is not possible to relate uniquely the $^{26}\text{Al}/^{27}\text{Al}$ ratios with other isotope ratios measured in these grains. For example, although the high ^{15}N excess in grain 4 would indicate a nova origin, the C and Si ratios cannot be fitted even approximately with model calculations in such objects²³. We hope that our correlated isotope measurements on single grains of stellar origin will stimulate new theoretical work on nucleosynthesis in their sources.

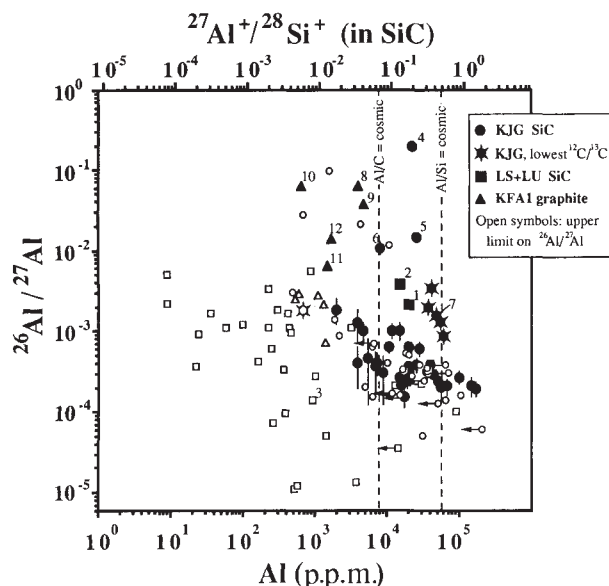


FIG. 2 $^{26}\text{Al}/^{27}\text{Al}$ ratios in Murchison SiC and graphite grains. In most of the LS-U (large) SiC grains only upper limits were obtained. In contrast, for at least 50% of the KJG (smaller) SiC grains and large graphite grains clear ^{26}Mg excesses are seen, with $^{26}\text{Al}/^{27}\text{Al}$ ratios ranging up to 0.2, 4,000 times the canonical value of 5×10^{-5} in refractory Ca-, Al-rich inclusions from primitive meteorites.

Received 27 August; accepted 9 November 1990.

1. Tang, M. & Anders, E. *Geochim. cosmochim. Acta* **52**, 1235-1244 (1988).
2. Amari, S., Anders, E., Virag, A. & Zinner, E. *Nature* **345**, 238-240 (1990).
3. Zinner, E., Tang, M. & Anders, E. *Geochim. cosmochim. Acta* **53**, 3273-3290 (1989).
4. Lewis, R. S., Amari, S. & Anders, E. *Nature* **348**, 293-298 (1990).
5. Wopenka, B. et al. *Meteoritics* **24**, 342 (1989).
6. Amari, S., Lewis, R. S. & Anders, E. *Lunar planet. Sci.* **21**, 19-20 (1990).
7. Amari, S., Zinner, E. & Lewis, R. S. *Meteoritics* (in the press).
8. Zinner, E., Wopenka, B., Amari, S. & Anders, E. *Lunar Planet. Sci.* **21**, 1379-1380 (1990).
9. Lattimer, J. M., Schramm, D. N. & Grossman, L. *Astrophys. J.* **219**, 230-247 (1978).
10. Catanzaro, E. J., Murphy, T. J., Garner, E. L. & Shields, W. R. *J. Res. Nat. Bur. Stand.* **70A**, 453-458 (1966).
11. Czyzak, S. J., Hirth, J. P. & Tabak, R. G. *Vistas Astr.* **25**, 337-382 (1982).
12. Nuth, J. A. *Nature* **318**, 166-168 (1985).
13. Frenklach, M., Carmer, C. S. & Feigelson, E. D. *Nature* **339**, 196-198 (1989).
14. Wasserburg, G. J. & Papanastassiou, D. A. in *Essays in Nuclear Astrophysics* (eds Barnes, C. A., Clayton, D. D. & Schramm, D. N.) 77-140 (Cambridge Univ. Press, 1982).
15. Clayton, D. D. *Nature* **257**, 36-37 (1975).
16. Tielens, A. G. G. M. in *Carbon in the Galaxy, Rep. NASA CP-3061* (eds Tarter, J. C., Chang, S. & DeFreese, D. J.) (NASA, Washington, DC, 1990).
17. Nørgaard, H. *Astrophys. J.* **236**, 895-898 (1980).
18. Cameron, A. G. W. *Icarus* **60**, 416-427 (1984).
19. Dearborn, D. S. P. & Blake, J. B. *Astrophys. J.* **288**, L21-L24 (1985).
20. Prantzos, N. & Cassé, M. *Astrophys. J.* **307**, 324-331 (1986).
21. Clayton, D. D. *Astrophys. J.* **280**, 144-149 (1984).
22. Arnould, M. in *Proc. Accelerated Radioactive Beams Workshop, Rep. TRI-85-1* (eds Buchmann, L. & D'Auria, J.) 29 (TRIUMF, 1985).
23. Woosley, S. E. & Hoffman, R. D. in *Nucleosynthesis and Chemical Evolution* (eds Hauck, B., Maeder, A. & Meynet, G.) 60-68 (Observatoire de Genève, 1986).
24. Arnett, W. D. & Wefel, J. P. *Astrophys. J.* **224**, L139-L142 (1978).

25. Truran, J. W. & Cameron, A. G. W. *Astrophys. J.* **219**, 226–229 (1978).
 26. Woosley, S. E. & Weaver, T. A. in *Nucleosynthesis and its Implications on Nuclear and Particle Physics* (eds Audouze, J. & Mathieu, N.) 145–166 (Reidel, Dordrecht, 1986).
 27. Gehr, R. D. in *Interstellar Dust* (eds Allamandola, L. J. & Tielens, A. G. G. M.) 445–453 (Kluwer, Dordrecht, 1989).

ACKNOWLEDGEMENTS. We thank Tang M. for the chemical preparation of HM and HN, A. Virag for the detailed characterization of the LU and LS grains, B. Wopenka for unpublished Raman data, J. Foote for grain mounting and maintenance of the ion probe and E. Inazaki for help with isotope measurements. S.A. and E.Z. thank R. Walker for his interest and support. This work was supported by NSF (E.Z.) and NASA (E.A.).

Biological control of surface temperature in the Arabian Sea

Shubha Sathyendranath^{*†}, Albert D. Gouveia[‡],
 Satish R. Shetye[‡], P. Ravindran^{*} & Trevor Platt[†]

^{*} Department of Oceanography, Dalhousie University, Halifax, Nova Scotia, Canada, B3H 4J1

[†] Biological Oceanography Division, Bedford Institute of Oceanography, Box 1006, Dartmouth, Nova Scotia, Canada, B2Y 4A2

[‡] Physical Oceanography Division, National Institute of Oceanography, Dona Paula, Goa 403 004, India

BY far the dominant variable parameter controlling the absorption cross-section for short-wavelength solar radiation incident on the ocean surface is the concentration of photosynthetic pigment contained in phytoplankton cells^{1,2}. The abundance of phytoplankton depends on the intensity of incident radiation and on the supply of essential nutrients (nitrogen in particular). A higher abundance increases absorption of radiation and thus enhances the rate of heating at the ocean surface. In the Arabian Sea, the southwest monsoon promotes seasonal upwelling of deep water, which supplies nutrients to the surface layer^{3,4} and leads to a marked increase in phytoplankton growth. Using remotely sensed data on ocean colour, we show here that the resulting distribution of phytoplankton exerts a controlling influence on the seasonal evolution of sea surface temperature. This results in a corresponding modification of ocean–atmosphere heat exchange on regional and seasonal scales. Thus we show that this biological mechanism may provide an important regulating influence on ocean–atmosphere interactions.

The thermal structure of the surface (mixed) layer of the ocean depends on the various processes acting in it and on it. Heat (sensible and latent) and momentum are exchanged between ocean and atmosphere by turbulent processes acting at the free surface and at the base of the mixed layer^{5,6}. In addition, radiative fluxes (input of short-wavelength radiation, long-wavelength back radiation) contribute to the net heat balance of the layer. In the ocean itself, attenuation of visible light depends on a suite of competing absorbers, of which the most important are the photosynthetic pigments contained in autotrophic microplankton^{1,2}. The concentration of chlorophyll-like pigments in the surface layer of the ocean varies seasonally and regionally by at least four orders of magnitude, modulating the total absorption coefficient.

An increased local absorption coefficient in the mixed layer results in heating being localized closer to the surface, rather than distributed through the water column⁷. This strengthens the vertical gradient of density, reducing the effectiveness of turbulence as an agent of vertical mixing, and thus decreases the thickness of the layer over which the total heating is distributed. Moreover, in times of nutrient sufficiency, stronger vertical stratification is more favourable to net phytoplankton growth⁸, reinforcing the initial response.

Despite the enormous dynamic range of chlorophyll concentrations in the sea, and the often dominant contribution of chlorophyll to absorption of visible light, models of ocean heating have, with few exceptions^{9,10}, avoided explicit treatment of the role of phytoplankton. Kirk⁷ considered a hypothetical

case in which the effect of water colour on solar heating was simulated by an increase of a factor of 12 in the absorption at 440 nm. He found a differential heating ($\sim 0.8^\circ\text{C d}^{-1}$) attributable to the influence of dissolved substances. Regional studies, however, are lacking, partly because of the absence of data on the chlorophyll fields at the required horizontal scales¹¹. Recently data on basin-scale chlorophyll fields have become available from remote sensing of ocean colour^{12,13}.

Here we discuss a model for heating the surface layer of the Arabian Sea, in which the concentration of chlorophyll is an important parameter, and deduce the amplitude of the variation in sea surface temperature that can be ascribed to variations in abundance of phytoplankton. The underlying model is an implementation¹¹ of the well-known Kraus–Turner model⁵. It is applied in 2° squares along a zonal strip at 10°N in the Arabian Sea, taking explicit consideration of the effect of chlorophyll concentration on the attenuation coefficient for short-wavelength radiation. We chose this particular region because its temperature and pigment fields are known to vary in space and time under the influence of the monsoons, and because of the special role that the tropics are believed to have in the coupling of the ocean with the atmosphere¹⁴. Calculations were carried out on the monthly averaged data for 1979.

We assume that 50% of the short-wavelength radiation incident on the ocean surface would be absorbed in the top 1 m of the water column⁷. The fate of the other (penetrative) part was calculated according to a model¹⁵ resolved with respect to wavelength, using the near-surface chlorophyll field derived from satellite (Coastal Zone Colour Scanner) data¹³ as an index of chlorophyll concentration in the mixed layer. We calculated the fraction F of the incident flux lost in the mixed layer by subtracting the flux passing downwards through the base of the mixed layer, integrated over wavelength, from the incident flux. The attenuation coefficient ν for the mixed layer is given by $\nu = (-\log_e F)/h$, where h is the depth of the mixed layer at the beginning of the month in question. The value of h , like the rest of the parameters except ν , is the climatological mean of the observational data^{16,17}. The surface temperature, T_s , and the temperature gradient at the base of the mixed layer $(\partial T/\partial z)_h$ were assumed known^{11,17} at the beginning of the month, as were the monthly mean values of the air–sea fluxes of momentum and heat¹⁸. The model was then integrated for one month to find the new values T_s^+ and h^+ .

Figure 1a, b shows the regional-scale response of the phytoplankton of the Arabian Sea to forcing by the monsoons. The difference between the January picture (northeast monsoon) and the September picture (southwest monsoon) is striking: the enhanced nutrient supply due to the upwelling results in a large and widespread increase in biomass. We used the monthly biomass for the zonal strip at 10°N (Fig. 1c) to estimate mean ν for each 2° square, and hence the heating rate. Figure 2 shows ΔT_s , the difference between sea surface temperature calculated from the data in Fig. 1c and that calculated assuming no contribution to ν from biological terms. Because the calculation is reset at the beginning of each month, ΔT_s ($^\circ\text{C month}^{-1}$) may be regarded as the biological contribution to the rate of heating of the mixed layer; ΔT_s can exceed $1^\circ\text{C month}^{-1}$. Although ΔT_s is highest near the coasts, the high values are not confined to the coast. The highest values occur on the eastern side of the region in March–April but on the western side in October. Moreover, ΔT_s is not just a simple mapping of ν or of the biomass field: because ν is not linear in the biomass, a positive change in biomass will not increase the heating rate by the same amount as a negative change of the same magnitude would reduce it (Fig. 3). Thus it is important to know the distribution of biomass about the mean, as well as the mean itself, for a fuller understanding of the biological contribution to surface heating rate.

Our conclusion, that variations in phytoplankton biomass in the Arabian Sea make a dominant contribution to the heating

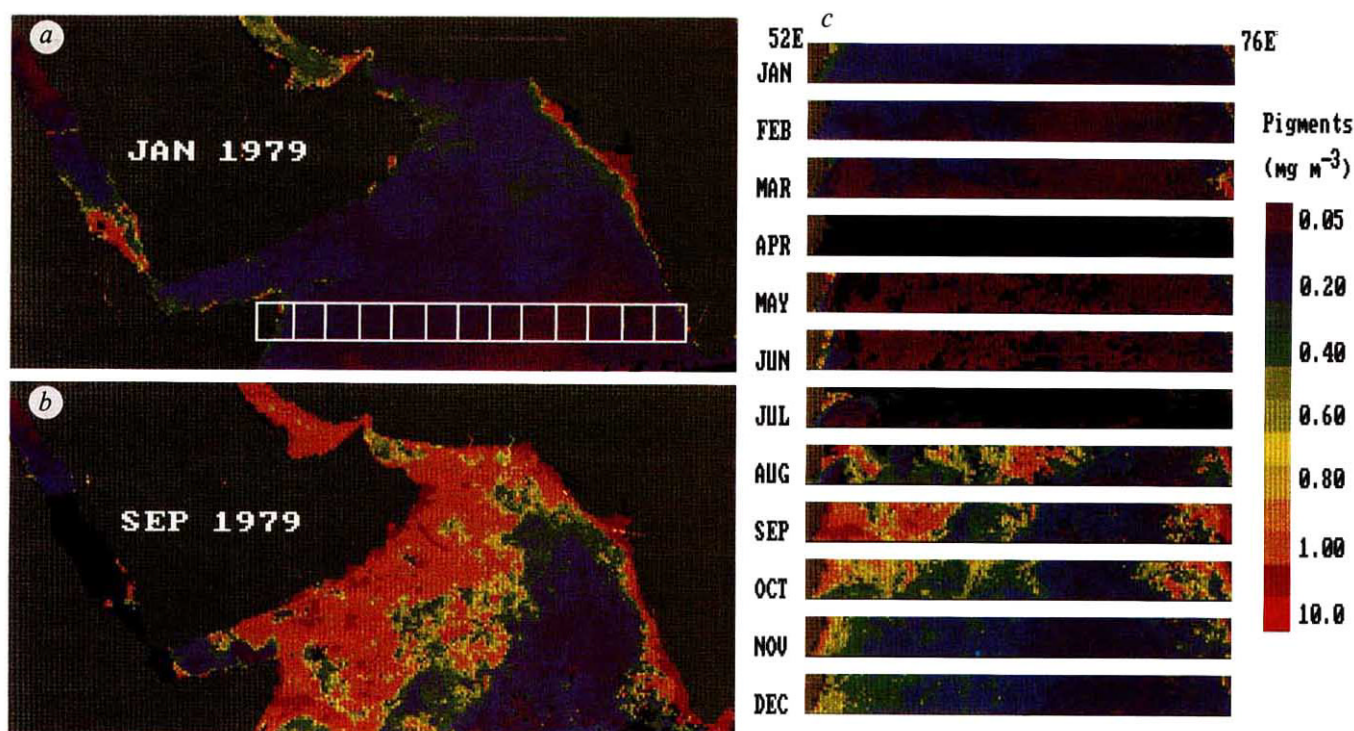


FIG. 1 Biomass fields for phytoplankton (chlorophyll-like pigments) in the Arabian Sea produced from ocean colour data collected by the Coastal Zone Colour Scanner in 1979. The evolution of the biomass field is forced by the timing and intensity of the monsoons. *a*, The situation in January under the influence of the northeast monsoon. The zonal strip shows the boxes for

which the heating rate was calculated for each month. *b*, The situation in September under the influence of the southwest monsoon. *c*, Mean biomass along a 2° zonal strip at 10° N in the Arabian Sea for each month of 1979. Black indicates missing data: in the calculations, missing data were replaced by averaging the data for adjacent months.

of the surface layer, is not dependent on our choice of mixed-layer model. In a comparison of various mixed-layer models, Martin¹⁹ found that all models examined were strongly sensitive to the optical attenuation coefficient, and that the more advanced turbulence closure models did not offer significant improvements over simpler integrated models such as we used. Variations in sea surface temperature on the spatial scales and timescales

characteristic of the phytoplankton distributions will lead inevitably to variations in the atmosphere on similar scales: the immediate response to a local increase in surface temperature will be an increase in the flux of moisture and an increase in the contributions of sensible heat, long-wavelength radiation and latent heat to the exchange of heat between the ocean and atmosphere^{20,21}. For the Arabian Sea, Kershaw²² has shown that knowledge of anomalies in sea surface temperature allows more accurate prediction of the evolution of atmospheric depressions associated with the monsoon. More generally, it is known that tropical storms develop preferentially over warmer waters²¹. These arguments justify the growing use of sea surface temperature as inputs to operational models for weather prediction in the tropics²¹. Our results indicate that specification of the local heating rate of the surface waters would be improved by consideration of the effect of the local biomass. The best tool to investigate the relevant spatial scales and timescales in the phytoplankton distributions will be the regional and seasonal chlorophyll fields as revealed by remote sensing of ocean colour.

The maximum rate of biological heating we calculate for the Arabian Sea is $\sim 4^\circ\text{C month}^{-1}$, in the period from August to September, an important contribution to the net heat flux. It may be compared with a maximum observed cooling rate (July) due to upwelling of $\sim 2.5^\circ\text{C month}^{-1}$ for the climatological mean^{11,17}. The effect of the presence of phytoplankton is asymmetrical: it diminishes the rate of cooling during the cooling (upwelling) season and enhances the rate of heating during the heating season. These results may be of more general significance. Because the biological effect is always positive, the surface layer of the world ocean is warmer than it would be in the absence of phytoplankton. Physical processes will communicate some of the surface heat gain to the deep ocean. Further, the heat trapped in the surface layer remains available for

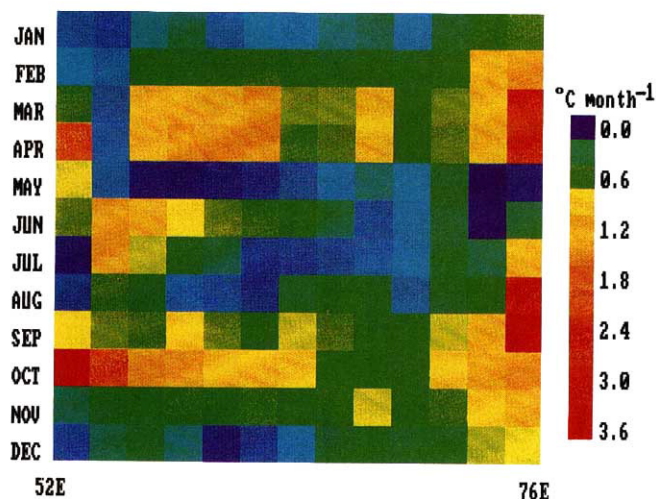


FIG. 2 Results of calculations (using the biomass fields in Fig. 1c) for ΔT_s , the biological contribution to rate of heating of the surface mixed layer, averaged over 2° squares at 10° N in the Arabian Sea for each month of 1979.

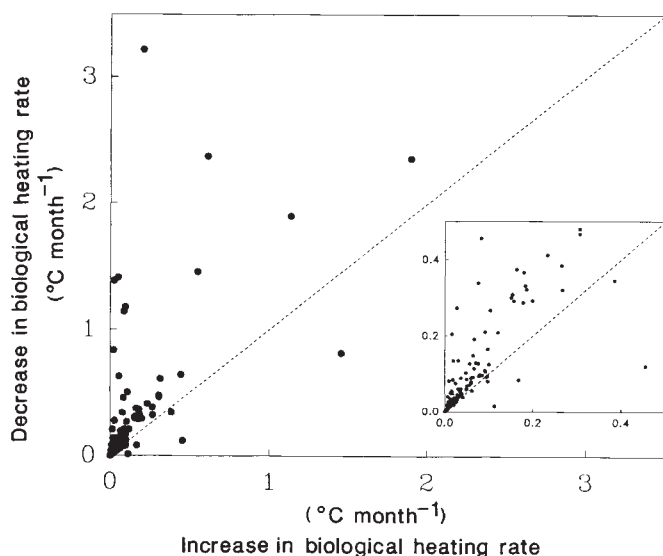


FIG. 3 Effect of local fluctuations in the biomass field on the biological component of heating rate for the mixed layer. Ordinate is the decrease in heating rate that would ensue if the mean biomass for any of the 2° squares in any month were decreased by an amount equal to the standard deviation of the biomass in the square in that month. The abscissa is the increase in heating rate that would be associated with an increase in mean biomass by one standard deviation. The inset shows an expanded view of the data near the origin. The broken lines have slopes of unity. The points are not evenly distributed about these lines. Because of the fundamental asymmetry due to the nonlinearity of the responsible processes, it is necessary to know the variance structure of the biomass field, as well as the large-scale features, for calculation of the biological contribution to the heating rate.

transfer back to the atmosphere. Notwithstanding these two moderating effects, the temperature contrast between surface and deep water is almost certainly enhanced by the presence of phytoplankton. Given that the north-south heat transport in the ocean is dominated by meridional overturning²³, a flux that depends directly on the temperature contrast between surface and deep waters, we can conclude that biological processes in the surface layer influence the net transport of heat by the ocean as a whole. □

Received 9 July; accepted 8 November 1990.

- Smith, R. C. & Baker, K. S. *Limnol. Oceanogr.* **23**, 247-259 (1978).
- Morel, A. *Boundary-Layer Met.* **18**, 177-201 (1980).
- Ryther, J. H. *et al. Limnol. Oceanogr.* **11**, 371-380 (1966).
- Sen Gupta, R. & Naqvi, S. W. A. *Deep Sea Res.* **31**, 671-706 (1984).
- Kraus, E. B. & Turner, J. S. *Tellus* **19**, 98-105 (1967).
- Denman, K. L. *J. phys. Oceanogr.* **3**, 173-183 (1973).
- Kirk, J. T. O. *J. geophys. Res.* **93**, 10897-10908 (1988).
- Sverdrup, H. U. *J. Cons.* **18**, 287-295 (1953).
- Lewis, M. R., Cullen, J. J. & Platt, T. *Mar. Ecol. Progr. Ser.* **15**, 141-149 (1984).
- Simonot, J.-Y., Dollinger, E. & Le Treut, H. *J. geophys. Res.* **93**, 8193-8202 (1988).
- Shetye, S. R. *J. mar. Res.* **44**, 521-542 (1986).
- Banse, K. & McClain, C. R. *Mar. Ecol. Progr. Ser.* **34**, 201-211 (1986).
- Feldman, G. *et al. Eos* **70**, 634-635; 640-641 (1989).
- Philander, S. G. & Rasmusson, E. M. *Adv. Geophys.* **28**, 197-215 (1985).
- Sathyendranath, S. & Platt, T. *J. geophys. Res.* **93**, 9270-9280 (1988).
- Levitus, S. *Climatological Atlas of the World Ocean*, NOAA Prof. Pap. No. 13 (U.S. Government Printing Office, Washington, DC, 1982).
- Hastenrath, S. & Lamb, P. *Climatic Atlas of the Indian Ocean Part I* (University of Wisconsin Press, 1979).
- Hastenrath, S. & Lamb, P. *Climatic Atlas of the Indian Ocean Part II* (University of Wisconsin Press, 1979).
- Martin, P. J. *J. geophys. Res.* **90**, 903-916 (1985).
- Haney, R. L. *J. phys. Oceanogr.* **1**, 241-248 (1971).
- Kershaw, R. Q. *J. R. met. Soc.* **114**, 325-345 (1988).
- Kershaw, R. *Nature* **315**, 561-563 (1985).
- Manabe, S. & Stouffer, R. J. *J. Clim.* **1**, 841-866 (1988).

ACKNOWLEDGEMENTS. We thank D. Wright for a review and for his suggestion regarding meridional overturning and oceanic heat transport. This work was supported by the U.S. Office of Naval Research, NASA, the Department of Fisheries and Oceans (Canada), and the Natural Sciences and Engineering Research Council of Canada. The collaboration between India and Canada was facilitated through the UNDP-Government of India Project on Remote Sensing.

The role of fluorine in carbonatite magma evolution

Bruce C. Jago & John Gittins

Department of Geology, University of Toronto, Toronto, Canada M5S 3B1

CARBONATITE magmas require some agent in solution to maintain them in the liquid state at geologically relevant temperatures and pressures¹. Although the liquidus temperatures of carbonate systems are greatly lowered in the presence of water², an unrealistically large water content (~95 wt%) is required for maximum lowering of the minimum melting temperature². Here we report experimental results which show that, in several carbonate systems, ~8 wt% fluorine lowers the minimum melting and liquidus temperatures to a similar extent as do these very large amounts of water. Thus, although water may well be present in most carbonatite magmas, it is neither the only nor necessarily the main agent by which they can remain liquid. Fluorine has the further effect of breaking the 'thermal barrier' imposed by the nyerereite composition in the system $\text{Na}_2\text{CO}_3\text{--K}_2\text{CO}_3\text{--CaCO}_3$, thereby allowing a low-alkali calcitic carbonatite magma to differentiate into a highly sodic carbonatite magma of the Oldoinyo Lengai type. Although this neither proves nor disproves a possible origin by liquid immiscibility, it restores the credibility of fractional crystallization as an important process in developing an alkali enrichment trend in carbonatites.

Strong evidence for the importance of fluorine in the evolution of carbonatite magma is provided by the compositions of carbonatites and their minerals³⁻⁵, and particularly by the unusually alkalic carbonatite lava flows at Oldoinyo Lengai which contain very little water (<0.5 wt% H_2O) but up to 15 wt% F+Cl in approximately equal amounts⁶. Because there is so little water in these rocks, it cannot be the agent responsible for their low liquidus temperatures of 500-590 °C (refs 7-9); F or

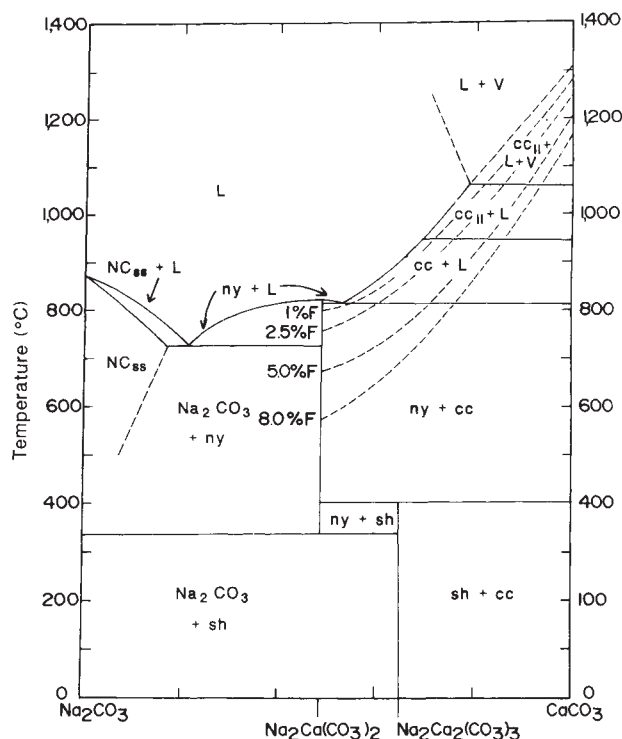


FIG. 1 Phase relations at 1 kbar in the system $\text{Na}_2\text{CO}_3\text{--CaCO}_3$ (wt%), showing the liquidus surfaces (solid line), and the projected calcite liquidus surfaces (dashed lines) with 1.0, 2.5, 5.0 and 8 wt% fluorine. Abbreviations are: NC_{ss} , Na-Ca carbonate solid solutions; ny, $\text{Na}_2\text{Ca}(\text{CO}_3)_2$ (nyerereite); cc, CaCO_3 (calcite); sh, $\text{Na}_2\text{Ca}_2(\text{CO}_3)_3$ (shortite); L, liquid; V, vapour; F, fluorine.

Cl seems the most likely alternative. Because of the overwhelming evidence for the presence of fluorine in carbonatite magmas, we have studied its effect on carbonate liquids and compared it with that of water. Some indication of the effect of F has been known since 1964 when it was shown that 36 wt% of fluorite (CaF_2) dissolves in calcitic liquid at 1 kbar and lowers the liquidus by $\sim 440^\circ\text{C}$ (ref. 10).

We have studied the effect of fluorine alone, without the complicating effect of H_2O , by determining the minimum melting temperatures at 1 kbar in several carbonate systems (Table 1). Reductions in the minimum melting temperatures are 234–290 $^\circ\text{C}$ for H_2O , 167–284 $^\circ\text{C}$ for CaF_2 , and 165–291 $^\circ\text{C}$ for F. These are broadly comparable but the powerful effect of fluorine is shown by the fact that reductions requiring at least 95% H_2O are achieved by only 8% F. Similarly, minimum melting temperatures in Na_2CO_3 and CaCO_3 (calcite) require $\sim 40\%$ of CaF_2 and $\text{Na}_2\text{Ca}(\text{CO}_3)_2$ (nyerereite) requires 20%, but all three require only $\sim 8\%$ F.

Figure 1 illustrates the way in which fluorine also lowers the liquidus temperatures in the system Na_2CO_3 – CaCO_3 . This is a projection of the calcite liquidus at 1, 2.5, 5.0 and 8 wt% F. At 8 wt% F, fluorite is stabilized and a calcite/fluorite boundary curve is developed.

Fluorine is therefore an efficient agent for lowering the minimum melting temperatures of carbonate systems relevant to carbonatite petrogenesis. Consequently, it seems likely that it is an important constituent of most carbonatite magmas. The effect of fluorine may also help answer some of the lingering questions about the origin of the Oldoinyo Lengai carbonatite magma and the evolution of more common, less alkalic carbonatite magmas.

The first experimental study¹¹ to deal with the Oldoinyo Lengai carbonatite lavas presented the phase diagram for Na_2CO_3 – K_2CO_3 – CaCO_3 at 1 kbar. It contains the two principal carbonate minerals of the natural lava—nyerereite ($\text{Na}, \text{K})_2\text{Ca}(\text{CO}_3)_2$ and gregoryite (a sodium-rich Na–K–Ca carbonate). The nyerereite composition in Fig. 1 and in the system Na_2CO_3 – K_2CO_3 – CaCO_3 constitutes a thermal barrier that prevents liquids whose bulk compositions are in the calcite stability field from differentiating to compositions any more sodic than the calcite–nyerereite cotectic. The system establishes the importance of alkalis in lowering the liquidus temperature of calcite-rich liquids but seems not to support the development of a highly sodic carbonatite magma by fractional crystallization of a calcitic parent magma. This system, however, is excessively simple in that it does not consider F and Cl, which are major constituents of the lavas, and so its applicability is severely limited. We have extended this system to Na_2CO_3 – CaCO_3 – CaF_2 (Fig. 2) and Na_2CO_3 – CaCO_3 –F (Fig. 3), both at 1 kbar.

The effect of CaF_2 and F is to break the thermal barrier caused by nyerereite. In Na_2CO_3 – CaCO_3 – CaF_2 (Fig. 2) the join $\text{Na}_2\text{Ca}(\text{CO}_3)_2$ – CaF_2 is penetrated by the calcite stability field which terminates at a reaction point where calcite reacts with

TABLE 1 Melting data for carbonate systems at 1 kbar

System	Minimum melting temp ($^\circ\text{C}$)	Reduction ($^\circ\text{C}$)	Ref.
CaCO_3	1,300*		2
CaCO_3 – H_2O	740 at 90% H_2O	580	2
CaCO_3 – CaF_2	880	420	10
CaCO_3 –F	1,100†	~ 200	6
Na_2CO_3	871		12
Na_2CO_3 – H_2O	630 at 97% H_2O	241	12
Na_2CO_3 – CaF_2	678	193	6
Na_2CO_3 –F	712 at 8% F	159	6
$\text{Na}_2\text{Ca}(\text{CO}_3)_2$	817		11
$\text{Na}_2\text{Ca}(\text{CO}_3)_2$ – H_2O	582 at 97% H_2O	289	6
$\text{Na}_2\text{Ca}(\text{CO}_3)_2$ – CaF_2	587	284	6
$\text{Na}_2\text{Ca}(\text{CO}_3)_2$ –F	580 at 8% F	291	6
Na_2CO_3 – $\text{Na}_2\text{Ca}(\text{CO}_3)_2$	725		11
Na_2CO_3 – $\text{Na}_2\text{Ca}(\text{CO}_3)_2$ – H_2O	490† at 88% H_2O	235	6
Na_2CO_3 – $\text{Na}_2\text{Ca}(\text{CO}_3)_2$ – CaF_2	558	167	6
Na_2CO_3 – $\text{Na}_2\text{Ca}(\text{CO}_3)_2$ –F	560	165	6
CaCO_3 – $\text{Na}_2\text{Ca}(\text{CO}_3)_2$	813		11
CaCO_3 – $\text{Na}_2\text{Ca}(\text{CO}_3)_2$ – H_2O	579† at 94% H_2O	234	6
CaCO_3 – $\text{Na}_2\text{Ca}(\text{CO}_3)_2$ – CaF_2	597	226	6
CaCO_3 – $\text{Na}_2\text{Ca}(\text{CO}_3)_2$ –F	580	216	6

* Dissociates at 1,040 $^\circ\text{C}$.

† Estimated.

liquid to generate nyerereite and fluorite. Crystallization is complete at a eutectic consisting of nyerereite, fluorite and sodium carbonate solid solution. The crystallization of CaF_2 requires a F concentration in the liquid of $\geq 11\%$, however, and although such values seem to be reached in the late stages of crystallization, a more common value is closer to 3%. Therefore, this system does not readily explain the early fractionation of a calcitic liquid.

More directly applicable is the system Na_2CO_3 – CaCO_3 –F (Fig. 3), which has two ternary reaction points and a eutectic. Calcite that crystallizes from liquid initially in the calcite stability field is later eliminated by reaction with liquid to generate $\text{Na}_2\text{Ca}(\text{CO}_3)_2$ and CaF_2 . $\text{Na}_2\text{Ca}(\text{CO}_3)_2$ is eventually eliminated by reaction to form NC_{ss} and CaF_2 (NC_{ss} denotes a solid solution of Na–Ca carbonate). A significant feature of this system, in contrast to Na_2CO_3 – CaCO_3 – CaF_2 , is the very substantial enlargement of the NC_{ss} stability field. Consequently, fractional crystallization can readily move a liquid from the calcite stability field to the $\text{Na}_2\text{Ca}(\text{CO}_3)_2$ – NC_{ss} cotectic and so permit co-crystallization of the principal rock-forming minerals of the Oldoinyo Lengai lavas, NC_{ss} being essentially the mineral gregoryite.

We note that fluorite is not a liquidus phase in the phenocryst-rich lavas but it is common as a groundmass mineral in lava that has been squeezed out of cracks on the surface of some

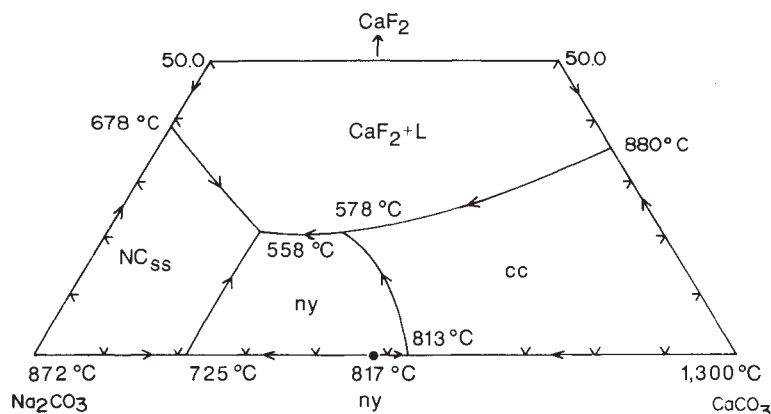
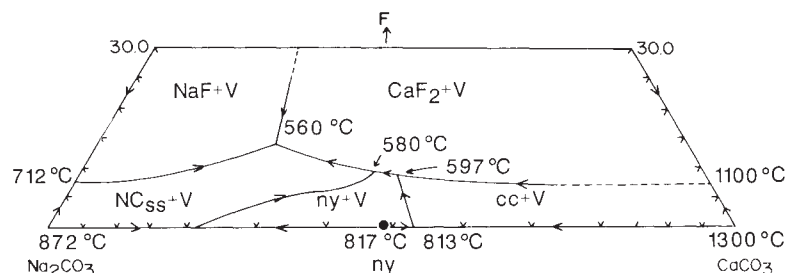


FIG. 2 Phase relations at 1 kbar in the carbonate-rich portion of the system Na_2CO_3 – CaCO_3 – CaF_2 (wt%).

FIG. 3 Phase relations at 1 kbar in the carbonate-rich portion of the pseudo-ternary system Na_2CO_3 – CaCO_3 –F (wt%).



flows. This lava represents the last liquid to crystallize; the rock is almost devoid of phenocrystic carbonate minerals but contains ~7.5% of both F and Cl. It is generated by the continuous crystallization of nyerereite and gregoryite, neither of which contains significant F. Consequently, F concentrates in the liquid to the level at which fluorite is stabilized. The composition of these squeezed-out liquids plots on the NC_{ss} – CaF_2 cotectic.

These F-bearing systems offer a mechanism whereby highly sodic carbonatite magmas of Oldoinyo Lengai type can develop through fractional crystallization of calcitic magmas (and presumably dolomitic magmas) that have a relatively low alkali content. The critical factor is that the parental magma must be dry because the separation of an aqueous fluid will remove alkalis and prevent their further accumulation¹.

It is also clear that, because of the large size of the calcite stability field in the presence of fluorine, fractional crystallization can develop the lesser degree of alkali enrichment that occurs in plutonic carbonatites where the alkalis are fixed in silicate minerals. Before the discovery of the effect of F, any removal of water and alkalis from the magma seemed to leave no means by which it could remain liquid at a geologically relevant temperature. Fluorine seems to offer a solution to this problem; it is known from mineralogical evidence to have been present in most carbonatite magmas and it now seems likely

that it had an important role in carbonatite magma evolution.

Although this work does not invalidate the concept of alkali-rich carbonatite magma being developed by liquid immiscibility it does demonstrate that fractional crystallization acting on calcitic, fluorine-bearing carbonatite magma with initially low alkali content remains a reasonable hypothesis in the continuing debate about the origins of the Oldoinyo Lengai carbonatite. □

Received 23 July; accepted 30 October 1990.

- Gittins, J. in *Carbonatites* (ed. Bell, K.) 580–600 (Unwin Hyman, London, 1989).
- Wyllie, P. J. & Tuttle, O. F. *J. Petrol.* **1**, 1–46 (1960).
- Dawson, J. B. in *Carbonatites* (eds Tuttle, O. F. & Gittins, J.) 155–168 (Wiley-Interscience, London, 1966).
- Dawson, J. B., Garson, M. S. & Roberts, B. *Geology* **15**, 765–768 (1987).
- Gittins, J. & McKie, D. C. *Lithos* **13**, 213–215 (1980).
- Jago, B. C. thesis, Univ. of Toronto (1990).
- Krafft, M. & Keller, J. *Science* **245**, 168–170 (1989).
- Dawson, J. B., Pinkerton, H., Norton, G. E. & Pyle, D. M. *Geology* **18**, 260–263 (1990).
- Keller, J. & Krafft, M. *J. Volcanol.* (in the press).
- Gittins, J. & Tuttle, O. F. *Am. J. Sci.* **262**, 66–75 (1964).
- Cooper, A. F. et al. *Am. J. Sci.* **275**, 534–60 (1975).
- Koster van Groos, A. F. *Am. Miner.* **75**, 667–675 (1990).

ACKNOWLEDGEMENTS. We thank the Natural Sciences and Engineering Research Council of Canada, Geological Survey of Canada, and the University of Toronto for support. J.G. also thanks the Warden and Fellows of Robinson College and the Department of Earth Sciences in the University of Cambridge for the tenure of a Bye Fellowship and sabbatical leave facilities.

Seismic quiescence at Parkfield due to detachment faulting

William D. Stuart

US Geological Survey, Pasadena, California 91106, USA

ON the San Andreas fault near Parkfield, California, rates of seismicity^{1,2} and geodetic line shortening^{3,4} have been lower since 1986 than before. Wyss *et al.*^{1,3} interpret the rate decreases as precursors to an imminent moderate earthquake (magnitude $M = 5.5$ – 6) and estimate the earthquake time to be $\text{March } 1991 \pm 1 \text{ yr}$. The earthquake was previously predicted to occur in $1988 \pm 5 \text{ yr}$ based on five of six similar earthquakes which occurred at intervals ~22 yr, the last one in 1966⁵. Here I present a model that attributes the rate changes to the initiation of slip on a subhorizontal detachment fault under the pending rupture area. In this model, the slip starts late in the seismic cycle, and diminishes the loading rate on the nearby San Andreas, thus acting as a buffer.

The proposed model is a simple variation of the generally accepted model for the stress on the San Andreas fault near Parkfield. In the current model the fault is a vertical plane divided into areas of locked fault, called patches, surrounded by areas of nearly steady fault slip (fault creep). Locked patches, the strongest parts of the fault, exist only on the upper 10–20 km of the fault, and earthquakes are generated when they are loaded to failure by the concentration of stress from adjacent fault creep. Fault creep is just the aseismic slippage of weak areas as

the Pacific and North America plates attempt sideways relative motion. Below 10–20 km the fault slips only by creep; shallower creep depends on the patch distribution. The $M = 8.3$ Fort Tejon earthquake in 1857 was due to the failure of a patch 10–20 km high, extending 350 km southeast from the Parkfield area (Fig. 1a). The 1966 mainshock was generated by the failure of an 8-km-wide patch centred ~6 km below Parkfield^{6,7}. Background seismicity is due to failure of smaller patches.

Figure 1b shows earthquake epicentres for magnitudes 2.0 and greater during the four-year period between 22 May 1986 and 21 May 1990. This period includes the quiescent period used in refs 1, 2 but ends five months later. The data are from the current US Geological Survey catalogue for Parkfield and differ slightly from the data used by Wyss *et al.*¹ and by Aviles and Valdes².

Figure 1c shows epicentres during the four-year period ending 21 May 1986. This four-year sample is representative of earlier seismicity distributions, and no other four-year sample between 1969 and 1986 has a lower rate. According to Wyss *et al.*¹ the occurrence rate of earthquakes of magnitude 2.4 and greater decreased by 80% in January 1986, and the rate for magnitude greater than 2.0 decreased by 46% in September 1986. In Fig. 1 the crustal volume for the rate calculations extends 5 km to either side of the fault and approximately between position XSC1 and Cholame. Aviles and Valdes² note that the seismic quiescence is most prominent northwest of the 1966 patch, between positions XSC1 and XMM1.

Two geodetic lines crossing the San Andreas fault near the 1966 patch show a decrease in the shortening rate of ~20% starting in late 1986³. The geodetic distance measurements are

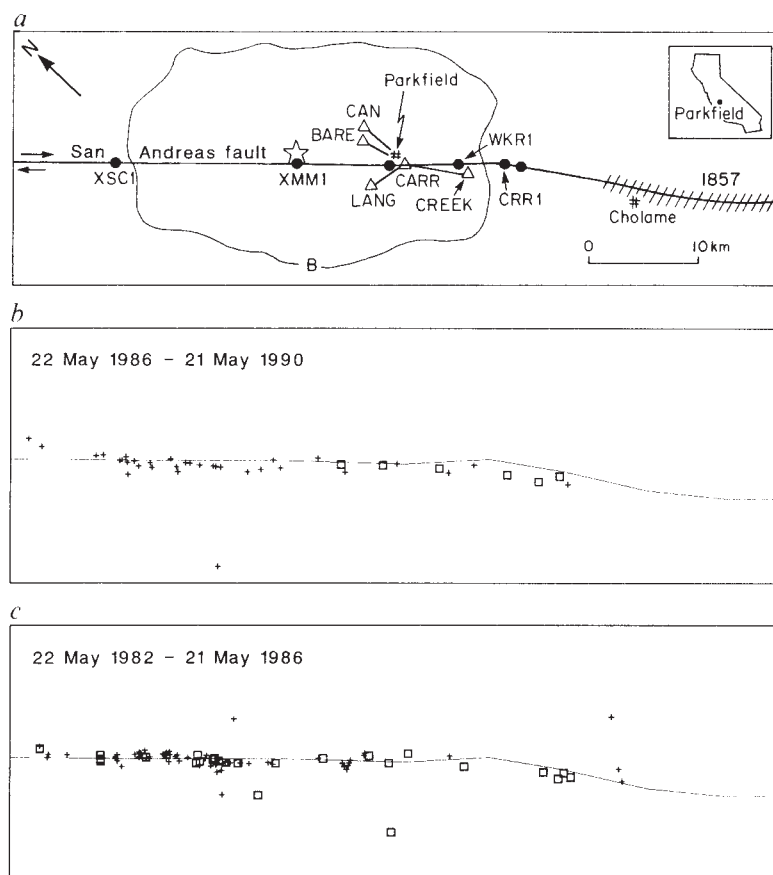


FIG. 1 *a*, Map of the San Andreas fault near Parkfield. Hachures mark the northwest tip of the section that slipped in the $M=8.3$ Fort Tejon earthquake. The star marks the epicentre of 1966 mainshock. Paired arrows show the direction of fault slip. Triangles mark benchmarks at the ends of geodetic lines; CARR is west of the fault. Dots show the locations of creep-meters where records were obtained before 1983. Wavy line B is a projection of the boundary of the postulated subhorizontal fault. (In ref. 3, line BARE was labelled BAR, and site CARR, CAR.) *b*, Earthquake epicentres during quiescence, 1986-1990. Only earthquakes within 10 km of the fault are plotted in *b* and *c*. 86% of all earthquakes in *b* and *c* together are shallower than 10 km. *c*, Earthquake epicentres before quiescence, 1982-1986. +, $2.0 \leq M \leq 2.5$; □, $M > 2.5$.

made by a two-colour-laser ranging instrument and have an uncertainty <1 mm (ref. 4). For line CAN (Fig. 1a) the estimated rates before 1 September 1986 are 9.59 ± 0.22 mm yr^{-1} and 10.49 ± 0.08 mm yr^{-1} , depending on treatment of a tare in the data (J. Langbein, personal communication). The corresponding rates after 1 September are 7.63 ± 0.11 mm yr^{-1} and 7.60 ± 0.06 mm yr^{-1} . The average of the two ratios of rates after and before 1 September is 0.76 ± 0.02 . By identifying noise sources Langbein (ref. 8 and personal communication) estimates that the actual uncertainty of the rate ratio for CAN is 0.15.

Line BARE (Fig. 1a) lies near CAN and shows a similar rate change in 1986. Its rate ratio is 0.800 ± 0.008 , but its actual uncertainty is also ~ 0.15 . Motion of benchmark CARR, the laser source, cannot be solely responsible for rate changes of CAN and BARE because the difference between BARE and LANG shows a comparable decrease in the rate³.

In my model, decreases in seismicity and geodetic rate are caused by initiation of slip on a postulated subhorizontal fault (B in Fig. 1a). There is no direct evidence of the fault, but it may be the westward extension of a detachment fault inferred to terminate in the fold belt associated with the $M=6.7$ Coalinga earthquake in 1983, 40 km north of Parkfield⁹. The same steady deep slip on the San Andreas fault that loaded the overhead patches involved in the 1857, 1966 and smaller earthquakes also loads a horizontal plane at B, and if the yield stress of fault B is a little less than the yield stress of the overhead San Andreas fault, fault B will slip first. But when fault B slips, the stressing rate on the fault above declines, and thus the creep rate around patches also decreases. Hence, the microseismicity rate should be lower because slower interpatch creep allows patches to fail less frequently. The rate decrease of geodetic lines is mainly due to slower fault slip near the ground surface.

To account for the observations, fault B must extend along the San Andreas approximately between positions XSC1 and

CARR in Fig. 1a. The southeast edge of B (edge normal to the San Andreas) may be in the vicinity of benchmark CREEK because this line shows a rate increase in 1986 instead of a decrease. The distance of the edges of B to the side of the San Andreas is estimated below from the observed rate change of line CAN.

The mechanical problem is three-dimensional, but here I use a two-dimensional approximation to calculate stress and displacement rates for comparison with observations. The two-dimensional model represents a slice of the Earth's crust normal to the fault in the quiescent zone. The model cross-section, Fig. 2a, shows the vertical San Andreas fault locked between depths d and D . This interval represents the part of the fault where patches cover most of the area. The top of the vertical fault, segment 4, extends from the surface to d , and slips at constant yield stress in response to slip on the bottom of the vertical fault, segment 1. Segment 4 is required because fault creep is observed at the ground surface at Parkfield (discussed below). Segment 1 extends from D to ∞ and slips at constant imposed velocity $\dot{u}_1$. Segment 4 is assumed to be weak enough that it starts sliding early in the earthquake cycle. Horizontal fault segments 2 and 3, $(0, D)$ to (H, D) and $(0, D)$ to $(-H, D)$ respectively, are assumed to have yield stresses sufficiently high that the faults start sliding late in the earthquake cycle. Symbols in Fig. 2a show the sense of relative displacement for the San Andreas fault. When the horizontal faults are active they transfer part of the stress concentration at the top of segment 1 out to $x = \pm H$ and away from geodetic lines and segment 4. For simplicity, fault slip on each segment is assumed to be uniform.

The fault model is broken into two separate models; the first is valid before the horizontal faults slip, the second after. In both models $D = 10$ km and $\dot{u}_1 = 30$ mm yr^{-1} . The first model is used to find d before the horizontal faults slip such that the theoretical shortening rate for CAN before 1986, $\dot{L}_b$, agrees with

the observed value of 10 mm yr^{-1} . Two equations must be satisfied

$$\dot{\tau}_1 + \dot{\tau}_4 = f_1 \dot{u}_1 + f_4 \dot{u}_4 = 0 \quad (1)$$

$$\dot{L}_b = g_1 \dot{u}_1 + g_4 \dot{u}_4 \quad (2)$$

Equation (1) expresses static equilibrium at the centre of segment 4 as it slides at constant yield stress. Equation (2) is the formula for lengthening rate of a geodetic line. $\dot{\tau}_1$ and $\dot{\tau}_4$ are rates of shear stress τ_{xz} due to slip of segments 1 and 4, respectively. Terms f_1 , f_4 , g_1 and g_4 describe the geometry; f_4 and g_4 contain d . Equations (1) and (2) are from stress and displacement solutions for infinite screw dislocations in an elastic half-space¹⁰. Each dislocation is at the boundary between slipping and locked segments of a fault, that is, at d and D . The resulting value for d , $\sim 5 \text{ km}$, is in rough agreement with that in refs 6 and 7 (see Table in Fig. 2b), but overestimates the depth to the locked section because the model does not account for loading by high fault slip rates northwest of Parkfield.

In the second model all four segments slip at the same time, and H is the parameter. Its value is determined by the rate change observed at CAN. Segments 2, 3 and 4 each obey an expression like equation (1), but with extra terms containing H for stresses due to slip on segments 2 and 3. At the centres of segments 2 and 3 the equations express equilibrium of shear stress rates $\dot{\tau}_{yz}$. Equation (2) is replaced by one including terms for segments 2 and 3. The equation for the ratio of the rates $\dot{L}_a/\dot{L}_b$ (where subscript a means after the rate change) is independent of $\dot{u}_1$ because $\dot{u}_2$, $\dot{u}_3$ and $\dot{u}_4$ are proportional to $\dot{u}_1$.

Figure 2b shows calculated $\dot{L}_a/\dot{L}_b$ as a function of H , with d as a parameter. The lower set of three curves applies when both the left and right horizontal faults are active; the upper set applies when just the right (east) fault, segment 2, is active. The observed range of $\dot{L}_a/\dot{L}_b = 0.76 \pm 0.15$ is shown by the vertical bar. When segments 2 and 3 both slip, the estimated value of H is between 3 and 12 km for $d = 5 \text{ km}$, and the corresponding slip rates on segments 2 and 3 are between 10.9 and 9.4 mm yr^{-1} . When segment 3 slip is prevented, the reduction of the rate along line CAN is less: the estimated value of H is essentially just greater than 5 km , and the slip rate of segment 2 is less than 14.1 mm yr^{-1} . Data from line BARE do not further constrain H

because the data uncertainties are similar and the lines are close together. There is no independent check of H , but the above values are plausible because they are less than the distance to the fold belt, $20\text{--}30 \text{ km}$.

The connection of model results with seismic quiescence is less certain than for geodetic lines because there is no precise mechanical theory relating fault slip or crustal deformation to a sequence of microearthquakes. At Parkfield the crustal volume producing microseismicity corresponds to a narrow volume enclosing the model fault $0 < y < D$, Fig. 2a, and the stress rate at position P is representative for that volume. With both segments 2 and 3 active, the calculated stress rate ratio $(\dot{\tau}_{xz})_a/(\dot{\tau}_{xz})_b$ at point P is between 0.69 and 0.48 for $d = 5 \text{ km}$. With segment 3 inactive the ratio is less than 0.71 . These values seem consistent with the seismicity rate decreases of 46% and 80% reported by Wyss *et al.*¹ Seismicity is not necessarily expected on a newly slipping detachment fault at Parkfield because the fault is deeper than most seismicity along the San Andreas fault. High temperature at depth may suppress earthquakes in both cases.

The model is less consistent with fault slip rates at the ground surface measured by creepmeters. Creepmeter data¹¹ are difficult to interpret because the San Andreas fault responded to stress perturbations of the Coalinga earthquake and the $M = 5.5$ North Kettleman Hills earthquake in August 1985, 17 km southeast of the Coalinga earthquake¹². For $d = 5 \text{ km}$ the calculated slip rate on segment 4 before any horizontal faults slip is 12.0 mm yr^{-1} . When segments 2 and 3 slip, segment 4 slips between 11.1 and 7.2 mm yr^{-1} for $3 < H < 12 \text{ km}$. When segment 3 does not slip the rate for segment 4 is less than 10.6 mm yr^{-1} . Creepmeter XCS1, which measured the highest rate of all creepmeters in Fig. 1a, shows slight deviation at any time from its overall rate of 22 mm yr^{-1} . The rate measured at XMM1 was 18 mm yr^{-1} before the Coalinga earthquake and 15 mm yr^{-1} after the North Kettleman Hills earthquake. At WKR1 the rate decreased from 8 to 6 mm yr^{-1} for the same time periods, but at CRR1 it increased from 2 to 6 mm yr^{-1} . Sites at which creepmeters measured increases in rate may mark the southeast edge of the horizontal fault. Rate changes were not measured at the other creepmeter sites in Fig. 1a.

The buffer fault model is a special case of fault interaction in which a secondary fault continuously perturbs the stress

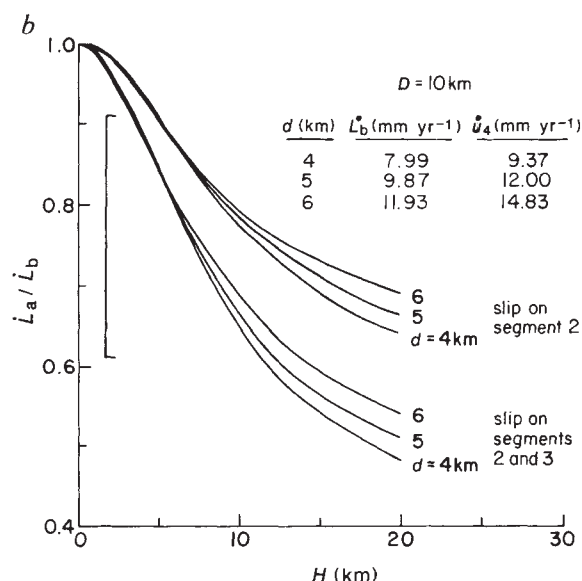
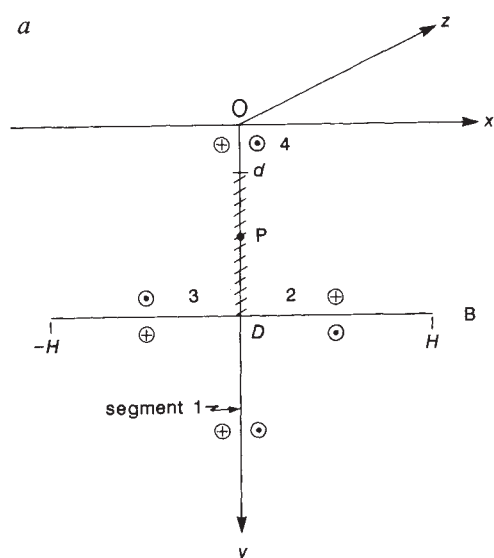


FIG. 2 a, Geometry of the model, looking northwest. x - z plane is the ground surface, y is down. Fault segment 1 has prescribed velocity $\dot{u}_1 = 30 \text{ mm yr}^{-1}$. Fault segments 2, 3, 4 slip at constant stress during the quiescent period. P is a point at the centre of the locked segment between d and D . Circled plus and dot indicate relative motion into and out of the page, respectively.

b, Theoretical ratio of shortening rates $\dot{L}_a/\dot{L}_b$ for line CAN (subscripts a and b mean rates after and before horizontal faults slip). Vertical bar shows the uncertainty on the ratio of the observed rates for CAN. Table shows shortening rates $\dot{L}_b$ for CAN and slip rates of segment 4, $\dot{u}_4$, for different values of d before segments 2 and 3 slip.

acting on the main fault. At Parkfield newly active vertical strike-slip faults could also decrease the loading rate on the San Andreas. One possible candidate is the Southwest Fracture Zone, a strike-slip fault sub-parallel to and 1–2 km southwest of the San Andreas near Parkfield, but the few displacement measurements available indicate that the slip rate is too low (R. Burford, personal communication). Slippage on a secondary fault might also increase the effective pressure on the patches of the San Andreas fault, strengthening them and delaying seismic failure; no such faults at Parkfield have been recognized yet. Alternatively, slip on a secondary fault might be impulsive instead of steady¹³. The best candidate at Parkfield is the North-Kettleman Hills earthquake, but the change in the coseismic stress calculated from dislocation theory is extensional and of right lateral sense where quiescence is strongest¹². □

Received 22 June; accepted 30 October 1990.

- Wyss, M., Bodin, P. & Habermann, R. E. *Nature* **345**, 426–428 (1990).
- Aviles, C. A. & Valdes, C. M. *Eos* **70**, 1228 (1989).
- Wyss, M., Slater, L. & Burford, R. O. *Nature* **345**, 428–431 (1990).
- Langbein, J. O., Burford, R. O. & Slater, L. E. *J. geophys. Res.* **95**, 2533–2552 (1990).
- Bakun, W. H. & Lindh, A. G. *Science* **229**, 619–624 (1985).
- Stuart, W. D., Archuleta, R. J. & Lindh, A. G. *J. geophys. Res.* **90**, 592–604 (1985).
- Harris, R. A. & Segall, P. *J. geophys. Res.* **92**, 7945–7962 (1987).
- Langbein, J. *Nature* (submitted).
- Wentworth, C. M. & Zoback, M. D. *Tectonics* **8**, 237–246 (1989).
- Rybicki, K. in *Physics and Evolution of the Earth's Interior* Vol. 3, 18–186 (Elsevier, New York, 1986).
- Schultz, S. S. *U.S. Geol. Surv. Open-File Rep.* 89–650 (1989).
- Simpson, R. W., Schultz, S. S., Dietz, L. D. & Burford, R. O. *Pure appl. Geophys.* **126**, 665–685 (1988).
- Mavko, G. M. *J. geophys. Res.* **87**, 7807–7816 (1982).

ACKNOWLEDGEMENTS. I thank J. Langbein for comments and calculations.

Localization of a human system for sustained attention by positron emission tomography

José V. Pardo^{*†}, Peter T. Fox^{†‡§}
& Marcus E. Raichle^{†‡§||}

Departments of ^{*}Psychiatry, [‡]Neurology and Neurological Surgery,
[§]Mallinckrodt Institute of Radiology, [†]McDonnell Center for Studies of
Higher Brain Function, Washington University Medical Center, St Louis,
Missouri 63110, USA

POSITRON emission tomographic (PET) studies of human attention have begun to dissect isolable components of this complex higher brain function, including a midline attentional system in a region of the anterior cingulate cortex^{1–3}. The right hemisphere may play a special part in human attention⁴; neglect, an important phenomenon associated with damage to attentional systems, is more severe, extensive and long-lasting after lesions to the right hemisphere. Here we use PET measurements of brain blood flow in healthy subjects to identify changes in regional brain activity during simple visual and somatosensory tasks of sustained attention or vigilance. We find localized increases in blood flow in the prefrontal and superior parietal cortex primarily in the right hemisphere, regardless of the modality or laterality of sensory input. The anterior cingulate was not activated during either task. These data localize the vigilance aspects of normal human attention to sensory stimuli, thereby clarifying the biology underlying asymmetries of attention to such stimuli that have been reported in clinical lesions.

Vigilance tasks require subjects to focus and sustain their attention to subtle sensory signals within a given modality, to minimize distractibility to irrelevant internal as well as external information, and to maintain adequate alertness for the duration of the study session. Participants report that these tasks involve

considerable effort, despite the scarcity of targets or minimal response output. Such a task paradigm has the potential advantage of activating neural systems relevant to sustained attention while minimizing activity in those brain regions involved in pre-attentive sensory processing, high-level processing selection of targets, and preparatory organization and production of response output.

Cerebral activation in healthy humans (23 subjects, 9 females and 14 males; mean age, 27 yr with s.d. of 5 yr, range 19–36 yr; 21 right-handed, one ambidextrous, and one left-handed) was measured as changes in regional cerebral blood flow using the intravenous H₂¹⁵O technique^{5,6} with the PETT VI system⁷. Two tasks requiring sustained attention to sensory input, one somatosensory task and one visual task, were used to isolate neural systems relevant to vigilance.

In the somatosensory vigilance task, volunteers focused their attention upon either their left or their right great toe as to detect brief pauses in a volley of suprathreshold touches. The great toe was touched with a von Frey hair which was applied manually at 3–5 Hz. The pauses lasted 1–3 seconds and were delivered (mean of 2.5 pauses per scan; s.d., 1.1 pauses; range of 0–5 pauses per scan) during the 40-second study interval (that is, about every 10 seconds). At the end of the study subjects were asked to report the number of pauses detected (mean of 4.2 pauses reported per scan; s.d., 2.5 pauses; range of 0–12 reported pauses per scan). To minimize possible confounding of neural activity related to eye movements, subjects were asked to maintain passive visual fixation on a dim central mark on a video screen while detecting pauses. The control state for this task was resting with eyes closed. To control for any effects of visual fixation, a further four subjects performed the same task (that is, they counted pauses to the right great toe while maintaining visual fixation), but in this case the control state was visual fixation. Individual image pairs (experimental state minus control state scans) were averaged together across subjects after anatomical normalization (stereotactic transformation)⁸ to generate a composite image of the average regional cerebral blood flow change associated with this task performance.

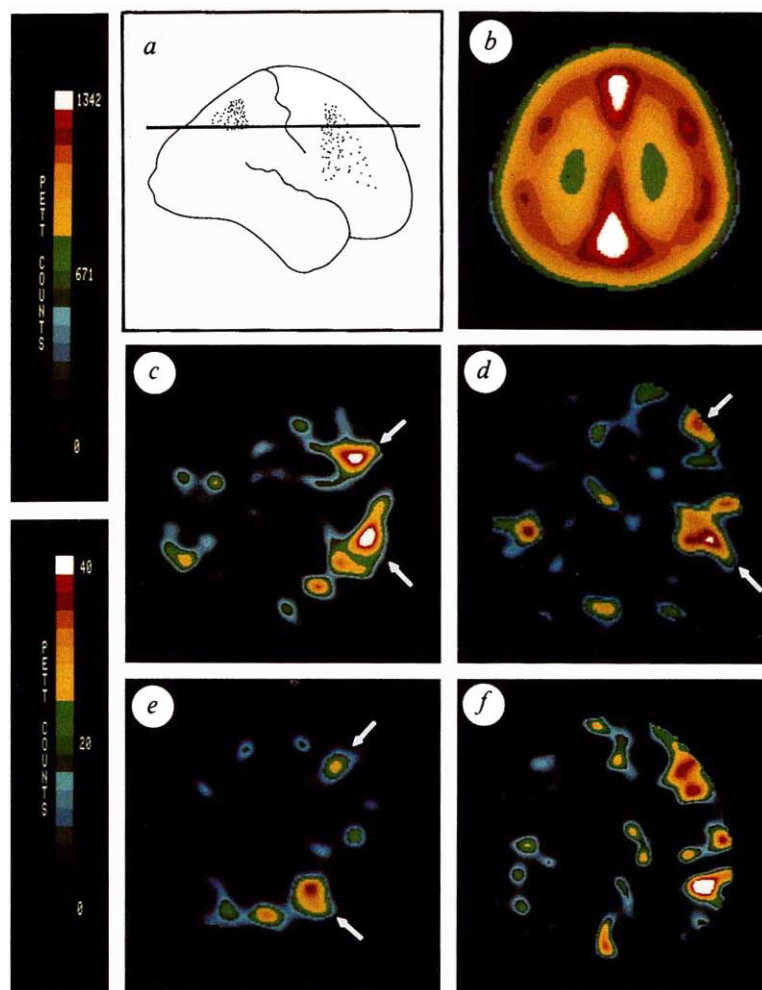
The averaged stereotactic images revealed not only the expected activation in contralateral somatosensory areas (not included in Table 1 because they were below the cutoff accepted for analysis (see Table 1 legend)) and occipital visual areas (see Table 1), but also cortical activation in regions previously not visualized during passive sensory stimulation^{9,10} (Figs 1, 2 and 3). Thus, monitoring the left great toe for pauses resulted in right prefrontal and right superior parietal responses (Fig. 1c; Fig. 3a). Monitoring the right great toe for pauses resulted also in right prefrontal and right superior parietal responses, with a smaller left superior parietal focus (Figs 1d and 3b). No activation of the left frontal cortex was evident, regardless of which great toe was stimulated; passive central fixation was clearly not responsible for the observed asymmetries (Fig. 1f). In this regard, another independent converging experiment comparing passive visual fixation with eyes closed at rest failed to activate this system (results not shown). Therefore the laterality of somatosensory input did not alter the laterality of the pre-dominant frontal-parietal activity.

In the visual vigilance task the same group of volunteers was asked to detect near-threshold luminance changes of a dim central fixation mark. Subjects were informed that threshold and subthreshold dimming might occur during the PET scan; they were to maintain fixation, monitor for dimming, and report the number of occurrences at the end of the study. No dimming occurred during the task, but most subjects reported some false alarms (mean of 5 false alarms per scan; s.d. 4; range of 0–15 false alarms). A stereotactic image of average regional cerebral blood flow change was generated from the visual vigilance task minus the eyes closed at rest scans⁸. As expected, medial striate responses were present¹⁰ and were greater than when subjects fixated on the same central fixation mark while attending to

|| To whom correspondence should be addressed.

FIG. 1 Stereotactic images of average cerebral activation induced by vigilance tasks; horizontal sections. *a*, Lateral view of the human brain (drawn to scale from the Talairach atlas¹¹) demonstrating section through horizontal plane at $Z=36$ mm (Z , vertical axis; see Table 1 legend) for sections *b-f* is shown. *b*, Tissue activity demonstrating resting blood flow. Orientation of all horizontal sections is anterior (top), posterior (bottom), left (to the left), right (to the right). *c*, Regional cerebral blood flow change between task (count pauses at left great toe and maintain passive visual fixation) and control (eyes closed at rest). Average of 18 image pairs from 10 subjects. *d*, Regional cerebral blood flow change between task (count pauses at right great toe and maintain passive visual fixation) and control (eyes closed at rest). Average of 17 image pairs from 9 subjects (different than those participating in *c*). *e*, Regional cerebral blood flow change between task (count dimming of central fixation mark) and control (eyes closed at rest). Average of 35 image pairs from 19 subjects (pooled from participants in *c* and *d*). *f*, Regional cerebral blood flow change between task (count pauses to right great toe and maintain passive visual fixation) and control (maintain passive visual fixation). Average of 10 image pairs from 4 subjects (different than the participants in *d*). The task and control states were selected to provide evidence that the frontoparietal activity is not related to passive visual fixation and to demonstrate replication of this activation during the somatosensory task with another group of subjects. Arrows demonstrate right prefrontal and superior parietal activation during vigilance. Upper scale (maximum 1,342) is for *b*; lower left scale (maximum 40) is for *c-f*.

METHODS. All subjects gave informed consent. Subject preparation was as described except that no arterial catheters were used¹⁰. As the change in local tissue radioactivity, (referred to as PETT counts in all figures) is linearly related to regional cerebral blood flow under the conditions of this study^{5,6}, the tissue activity changes are used synonymously with regional blood flow changes¹⁵. Each subject underwent multiple scans ~10 min apart. All tasks were presented in a counter-balanced manner. Differences between scans were calculated using pixel-by-pixel subtraction after linear scaling for changes in the global average activity of the brain (arbitrarily set to 1,000 PETT counts)^{8,15}. Anatomical normalization was performed by linear scaling to brain HD6 of the Talairach atlas¹¹. PET scan stereotactic averaging across subjects was described and validated previously⁸. Response locations were determined with a centre of mass algorithm using a 14-mm diameter sphere as the region of interest; the response magnitudes were calculated as the average activity within this sphere¹⁶. Approximate cytoarchitectonic fields



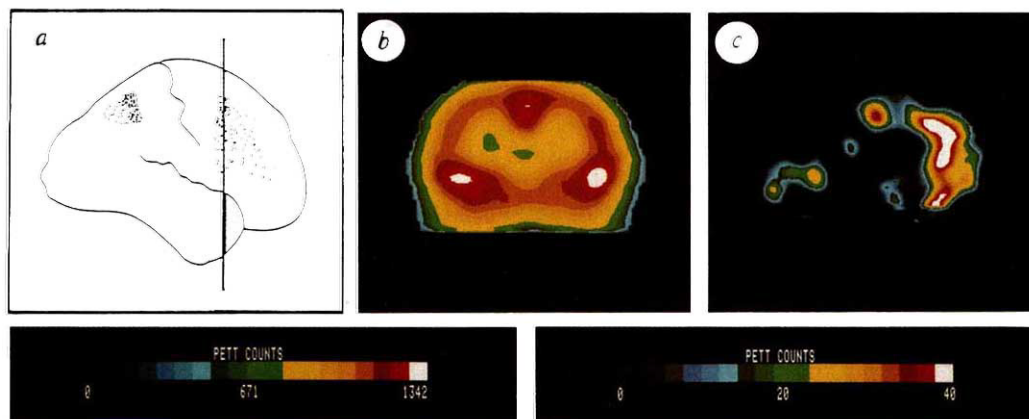
for responses were estimated from plates XLIV and LIX of the Talairach atlas¹¹.

their toes (Table 1, occipital). This finding is consistent with previous work on visual attention³. The right prefrontal and right superior parietal cortices were reproducibly and predominantly activated even with the central placement of the fixation mark (Figs 1e and 3c). In summary, the right prefrontal and superior parietal cortices were active during sensory pro-

cessing across two modalities, and the right-laterality of this system was not related to the laterality of sensory input.

The pattern of right prefrontal and right superior parietal activation during both vigilance tasks consisted of clusters of foci (Fig. 3 and Table 1). The stereotactic locations of all the responses, including those shown in Fig. 3, are given in

FIG. 2 Stereotactic images of average cerebral activation induced by vigilance tasks; coronal sections. *a*, Lateral view of the human brain demonstrating parietal and frontal regions activated. Coronal plane at $Y=26$ mm (Y , anteroposterior axis; see Table 1 legend) for sections *b* and *c* cuts through prefrontal activity in Fig. 1c. *b*, Tissue activity demonstrating resting blood flow. Orientation of coronal sections is superior (top), inferior (bottom), left (to the left), right (to the right). *c*, Regional cerebral blood flow change between task (count pauses at left great toe and maintain passive visual fixation) and control (eyes closed, rest). Note broad coronal band of activity along the right dorsolateral prefrontal cortex. Scale



at left (maximum 1,342) is for *b*; scale at right (maximum 40) is for *c*.

TABLE 1 Response coordinates and magnitudes

<i>a</i> Count pauses left toe (<i>n</i> =18)				
Region	<i>z</i>	<i>x</i>	<i>y</i>	Magnitude
1 Right occipital	6	-23	-68	56
2 Right occipital	44	-51	-23	41
3 Left occipital	4	27	-67	40
4 Right frontal	26	-43	27	38
5 Right frontal	30	-45	27	37
6 Right frontal	38	-43	25	36
7 Right temporal	10	-53	-29	34
8 Right frontal	4	-43	23	32
9 Right frontal	51	-24	17	30
10 Right temporal	20	-57	-25	30
11 Right frontal	52	-21	15	29
12 Right frontal	24	-21	39	29
13 Right frontal	48	-5	27	29
14 Left thalamus	14	5	-7	25
15 Right frontal	46	-9	17	24
16 Right parietal	48	-48	9	24
17 Right frontal	38	-24	-47	24

<i>b</i> Count pauses right toe (<i>n</i> =17)				
Region	<i>z</i>	<i>x</i>	<i>y</i>	Magnitude
1 Left occipital	8	29	-65	89
2 Right occipital	8	-27	-65	83
3 Right parietal	34	-51	-19	34
4 Right parietal	42	-49	-23	33
5 Left temporal	18	55	-19	33
6 Right frontal	31	-45	41	30
7 Right parietal	43	-49	-13	30
8 Left parietal	38	49	-17	30
9 Right frontal	28	-51	33	30
10 Right temporal	-6	-61	-17	30
11 Right frontal	31	-48	37	29

<i>c</i> Count dimming (<i>n</i> =35)				
Region	<i>z</i>	<i>x</i>	<i>y</i>	Magnitude
1 Left occipital	4	29	-64	119
2 Right occipital	6	-27	-65	110
3 Right parietal	34	-29	-51	37
4 Midline cerebellum	-6	5	-49	37
5 Left occipital	0	5	-55	35
6 Right parietal	36	-5	-61	31
7 Right parietal	48	-35	-35	31
8 Right frontal	34	-45	21	30
9 Right parietal	46	-39	-27	29
10 Right parietal	46	-49	-25	26
11 Right frontal	44	-31	17	25

All averaged images had statistically significant ($P < 0.01$) response outliers as assessed with the gamma-2 statistic⁸. From the distribution of all positive responses we report only those with a *z* score greater than 2.1, based on the standard deviation of the sample distribution, as previously described⁸. In this table *z* is the vertical axis (positive, superior); *x* is the right (negative)/left (positive) axis; *y* is the anterior (positive)/posterior (negative) axis¹⁰. Responses are numbered sequentially according to magnitude as normalized PET counts¹⁵ observed for each vigilance task. The location of those responses on the surface of the cerebral hemispheres is illustrated in Fig. 3.

Table 1. The parietal responses were concentrated (according to the stereotactic coordinates and Talairach's¹¹ estimates of Brodmann cytoarchitectonics) in areas corresponding to Brodmann area 7. The frontal responses were more dispersed and consisted of a variable coronal band of activity along the right dorsolateral convexity, corresponding to Brodmann areas 8, 9, 44 and 46 (with concentration in area 9, see Fig. 3). The concentration of major responses in the frontal and parietal regions is obvious across both vigilance tasks. In the somatosensory vigilance tasks, there seems also to be a temporal response which may be an additional component of the right-lateralized vigilance system. In no case was the anterior cingulate activated over the control reference state.

The predominant right frontal and parietal responses demon-

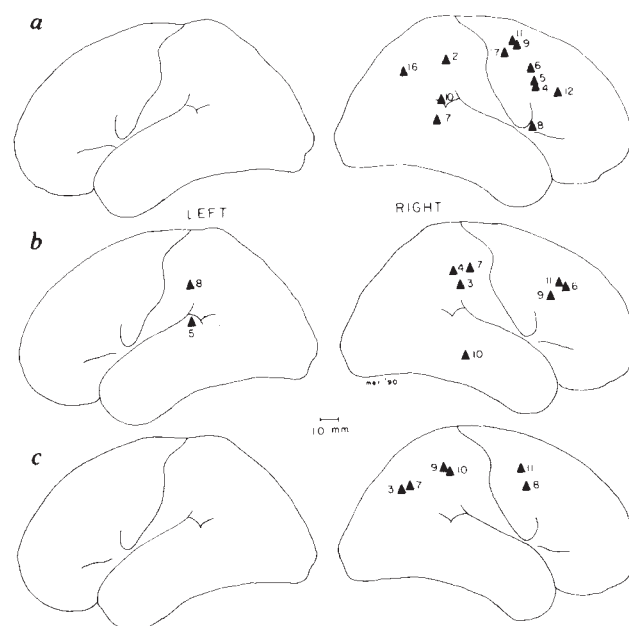


FIG. 3 Major responses ($Z > 2.1$) with stereotactic coordinates and magnitudes observed over the surface of both cerebral hemispheres during somatosensory and visual vigilance tasks. Responses are numbered in order of their relative magnitude. Exact magnitudes and stereotactic locations are given in Table 1. *a*, Responses during task (count pauses to left great toe while maintaining visual fixation) as compared with reference state (eyes closed at rest). *b*, Responses during task (count pauses to right great toe while maintaining visual fixation) as compared with reference state (eyes closed at rest). *c*, Responses during task (count dimming of central fixation mark) and reference state (eyes closed at rest).

strated in higher-order association cortices during vigilant behaviour across modalities defines anatomically a neural network involved in sustained attention to sensory input. The absence of anterior cingulate activation is consistent with the failure to recruit high-level processing selection systems necessary for the analysis of complex targets occurring at high rates^{1,2}. In contrast, the neural system activated during these vigilance tasks must relate to the 'on-line' analysis of the stimuli for relevant target properties. Work on the monkey has identified both somatosensory and visual cortical fields specialized for attentive analyses of sensory features, but these systems seem to process the contralateral visuospatial or somatospatial space^{12,13}. Whether both monkeys and humans have homologous neural circuits concerned with sustained attention is unknown. It is interesting that the prefrontal and parietal components of the mnemonic circuitry in the monkey are organized in parallel; also that the parietal regions associated with vision (Brodmann 7a) are posterior to the parietal regions associated with somatic sensation (Brodmann 7b)¹⁴.

These data identify and localize in healthy humans an anatomically distinct and asymmetric neural system mediating sustained attention to sensory stimuli. Vigilance, then, is one cognitive component of human attention that is right-lateralized. The vigilance system encompasses both right prefrontal and right superior parietal cortices and can operate independently of midline anterior attentional systems. These findings offer insight into the phenomenon, inferred from the literature on human lesions, of the right hemispheric dominance of human attention. □

Received 30 July; accepted 23 October 1990.

- Posner, M. I., Petersen, S. E., Fox, P. T. & Raichle, M. E. *Science* **240**, 1627-1631 (1988).
- Pardo, J. V., Pardo, P. J., Janer, K. W. & Raichle, M. E. *Proc. natn. Acad. Sci. U.S.A.* **87**, 256-259 (1990).
- Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L. & Petersen, S. E. *Science* **248**, 1556-1559 (1990).

4. Heilman, K. M., Watson, R. T., Valenstein, E. & Goldberg, M. E. in *Handbook of Physiology: Higher Cortical Functions of the Brain* (eds Plum, F. & Mountcastle, V. B.) **5**, 461–481 (Am. Physiol. Soc., Washington DC, 1987).
5. Herscovitch, P., Markham, J. & Raichle, M. E. *J. nucl. Med.* **24**, 782–789 (1983).
6. Raichle, M. E. et al. *J. nucl. Med.* **24**, 790–798 (1983).
7. Yamamoto, M., Ficke, D. C. & Ter-Pogossian, M. M. *IEEE Trans. nucl. Sci.* **29**, 529–533 (1982).
8. Fox, P. T., Mintun, M. A., Reiman, E. M. & Raichle, M. E. *J. cereb. Blood Flow Metab.* **8**, 642–653 (1988).
9. Fox, P. T., Burton, H. & Raichle, M. E. *J. Neurosurg.* **67**, 34–43 (1987).
10. Fox, P. T. et al. *J. Neurosci.* **7**, 913–922 (1987).
11. Talairach, J. & Szikla, G. et al. *Atlas d'Anatomie Stéréotaxique du Télencéphale* (Mason, Paris, 1967).
12. Goldman-Rakic, P. S. in *Handbook of Physiology: Higher Cortical Function of the Brain* (eds Plum, F. & Mountcastle, V. B.) **5**, 373–417 (Am. Physiol. Soc., Washington DC, 1987).
13. Fuster, J. M. *The Prefrontal Cortex* (Raven, New York, 1989).
14. Cavada, C., Goldman-Rakic, P. S. *J. comp. Neurol.* **287**, 393–421 (1989).
15. Fox, P. T. & Mintun, M. A. *J. nucl. Med.* **30**, 141–149 (1989).
16. Mintun, M. A., Fox, P. T. & Raichle, M. E. *J. cereb. Blood Flow Metab.* **9**, 96–103 (1989).

ACKNOWLEDGEMENTS. We thank the experimental subjects for their patience and efforts, P. J. Pardo for assistance and M. I. Posner for discussion. This research was supported by grants from the NIH, the McDonnell Center for Studies of Higher Brain Function, and NARSAD. J.V.P. is a Pfizer Fellow.

First skulls of the Early Eocene primate *Shoshonius cooperi* and the anthropoid-tarsier dichotomy

K. Christopher Beard*, Leonard Krishtalka* & Richard K. Stucky†

* Section of Vertebrate Paleontology, Carnegie Museum of Natural History, 4400 Forbes Avenue, Pittsburgh, Pennsylvania 15213, USA

† Department of Earth Sciences, Denver Museum of Natural History, Denver, Colorado 80205, USA

THE phylogenetic relationships of living tarsiers and extinct omomyid primates are critical for deciphering the origin and relationships of primate higher taxa, particularly anthropoids^{1–6}. Three competing phylogenetic hypotheses are: (1) tarsiers are most closely related to early Cenozoic Omomyidae^{5–8}, particularly genera such as *Necrolemur* from the late Eocene of Europe^{9–11}; (2) tarsiers share a more recent common ancestry with anthropoids than they do with any known omomyid^{2–4,12,13}; (3) tarsiers and/or omomyids are most closely related to strepsirhines¹⁴. The anatomy of four skulls of the early Eocene omomyid *Shoshonius cooperi*—the first cranial material recovered for this genus—strongly suggests that *Shoshonius* shares a more recent common ancestry with *Tarsius* than do either anthropoids or other Eocene omomyids for which cranial anatomy is known. If the primate suborder Haplorhini (anthropoids, omomyids, tarsiids) is monophyletic, the phylogenetic position of *Shoshonius* requires that anthropoids and *Tarsius* diverged by at least the early Eocene, some 15 million years before the first appearance of anthropoids in the fossil record^{15–17}.

The first known skulls of *Shoshonius cooperi* (CM 60492–60495) were recovered during the 1984–1987 field seasons from Carnegie Museum's Buck Spring Quarries 1 and 6 in the Lost Cabin Member, Wind River Formation, northeastern Wind River Basin, Wyoming¹⁸. The associated mammalian fauna (65 species) indicates a Lostcabinian age (latest Wasatchian or late early Eocene, about 50.5 Ma)^{18,19}.

The orbits of *Shoshonius* (Fig. 1a; diameter, 12.5–12.8 mm) are derived in being larger relative to skull length (27.8–29.2 mm) than in the Eocene omomyids *Necrolemur*, *Rooneyia* and *Tetonius*; however, they are smaller than in *Tarsius*^{7–9,11,12} (Fig. 3). Although none of the skulls of *Shoshonius* preserves a complete premaxilla, the shape of the dental arcade and preserved nasal region (Fig. 1a, 2) imply that the snout was greatly reduced compared with *Rooneyia* and *Necrolemur*, a derived condition shared with *Tarsius*. *Shoshonius* lacks a postorbital septum, in contrast to *Tarsius*, in which a partial postorbital septum is formed by periorbital bony flanges derived from the maxilla, squamosal, alisphenoid, and frontal¹². In this respect,

Shoshonius is primitive, resembling living and fossil strepsirhines and the omomyids *Tetonius*, *Necrolemur* and *Rooneyia*¹².

The auditory bullae of *Shoshonius* are highly inflated, particularly anteromedially, where they encroach upon the pterygoid region (Fig. 1b, 2). The lateral pterygoid wings attach to the bullae anterolaterally, a derived feature also present in some Eocene omomyids and *Tarsius*, but unlike the condition in *Rooneyia* and other primates¹⁰. As in *Tarsius* and *Rooneyia*, but in contrast to *Necrolemur* (and possibly *Tetonius*)⁷, there is no pneumatization of the petromastoid region in *Shoshonius*. A well developed basioccipital flange overlaps the bullar wall posteromedially, a derived feature unique to *Shoshonius* and *Tarsius* among living and fossil primates for which the relevant anatomy is known (including *Necrolemur* and *Rooneyia*)^{4,10,20}. This flange occurs more anteriorly on the bulla in *Tarsius*.

An intrabullar ectotympanic ring is preserved in place in CM 60494. The ring is joined to the internal face of the lateral bullar wall by a narrow, spool-shaped, multiseptate annular bridge similar to that in *Necrolemur*^{7,9}. The annular bridge is wider in *Rooneyia*^{7,21} and is absent in *Tarsius*, lorissoids and anthropoids⁴. In contrast to the condition in *Tarsius* and living catarrhines, the ectotympanic does not form an external tube in *Shoshonius*. Rather, the bone (either ectotympanic or petrosal) forming the anterior and posterior edges of the external acoustic meatus in *Shoshonius* exhibits slight lateral flaring (Fig. 1b). *Necrolemur* and *Rooneyia* have tubular structures of uncertain homology lateral to their bullae; they may be composed of ectotympanic^{7–9,21} or petrosal²².

The posterior carotid foramen (PCF) in *Shoshonius* pierces the bulla ventrolaterally, just posterior to the external acoustic meatus. This location, derived among primates, most closely approximates the PCF position in *Tarsius*, which is also ventrolateral but just anterior to the external acoustic meatus^{4,7,10,11,21}. The internal carotid canal in *Shoshonius* runs dorsally within a bony septum as in *Tarsius* and anthropoids, and crosses the anteroventral surface of the promontorium as the promontory canal. The presence or absence of a stapedia canal cannot be ascertained. It is not visible in any of the *Shoshonius* specimens at present.

Two longitudinal intrabullar bony septa occur in *Shoshonius*. The smaller, posterior one connects the internal carotid canal to the posterior bullar wall. The larger, anterior septum runs from the promontorium to the anterior end of the bulla. Thus, the intrabullar space is divided into medial and lateral cavities, except for a large hiatus between the septa just anterior to the promontorium. A single, more nearly complete intrabullar bony septum occurs in *Tarsius* and anthropoids, but the homology of the septum in these taxa is a contentious issue^{2–4,11}.

Shoshonius, like *Tarsius*, has a suprameatal foramen (SMF), albeit of smaller size. This derived feature is not known in any other primate (except possibly *Necrolemur*⁴); in *Tarsius* it is the site of an anastomosis between the posterior auricular artery and the ramus superior of the stapedia artery⁴. Because the stapedia artery is greatly reduced or absent in *Tarsius*⁴, this anastomosis is important in allowing the posterior auricular artery to irrigate the dura.

Dental similarities have long been used to ally omomyids and *Tarsius* phylogenetically^{7,8}. *Shoshonius cooperi* possesses derived dental features (for example, upper molar mesostyle and protocone fold; lower molar metastylid) that are absent in *Tarsius*⁷. Other related Eocene omomyines such as *Omomys*, *Chumashius* and *Loveina minuta* lack these dental specializations⁷ and are remarkably similar in molar morphology to *Tarsius*, but the skulls of these taxa have never been recovered. *Afrotarsius*, known only from a lower jaw from the early Oligocene of the Fayum, Egypt, may be a fossil tarsiid²³ or a primitive anthropoid^{24,25}.

In conclusion, *Shoshonius* and *Tarsius* share three cranial traits that are uniquely derived among primates for which the relevant anatomy is known, including: (1) a basioccipital flange

FIG. 1 Stereophotographs of two skulls of *Shoshonius cooperi*. *a*, Dorsal view of CM 60493. *b*, Ventrolateral view of CM 60494. The head of the malleus is protruding through the right external acoustic meatus. The left bulla, pictured here as complete, has since been prepared to reveal the internal anatomy reconstructed in Fig. 2 and described in the text.

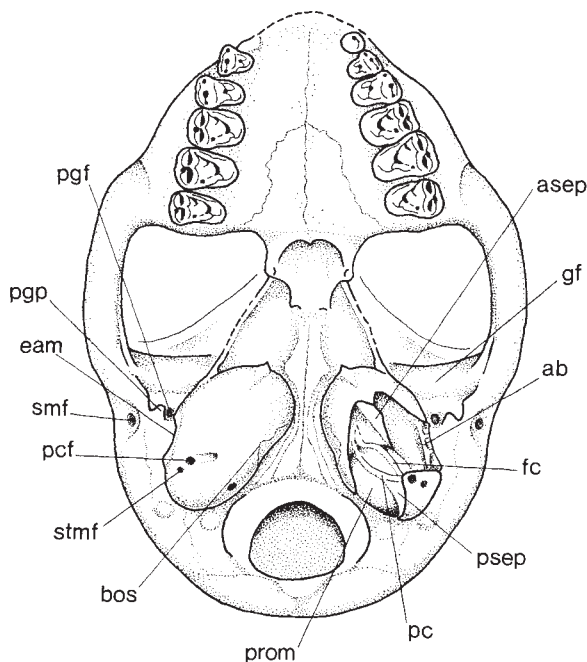
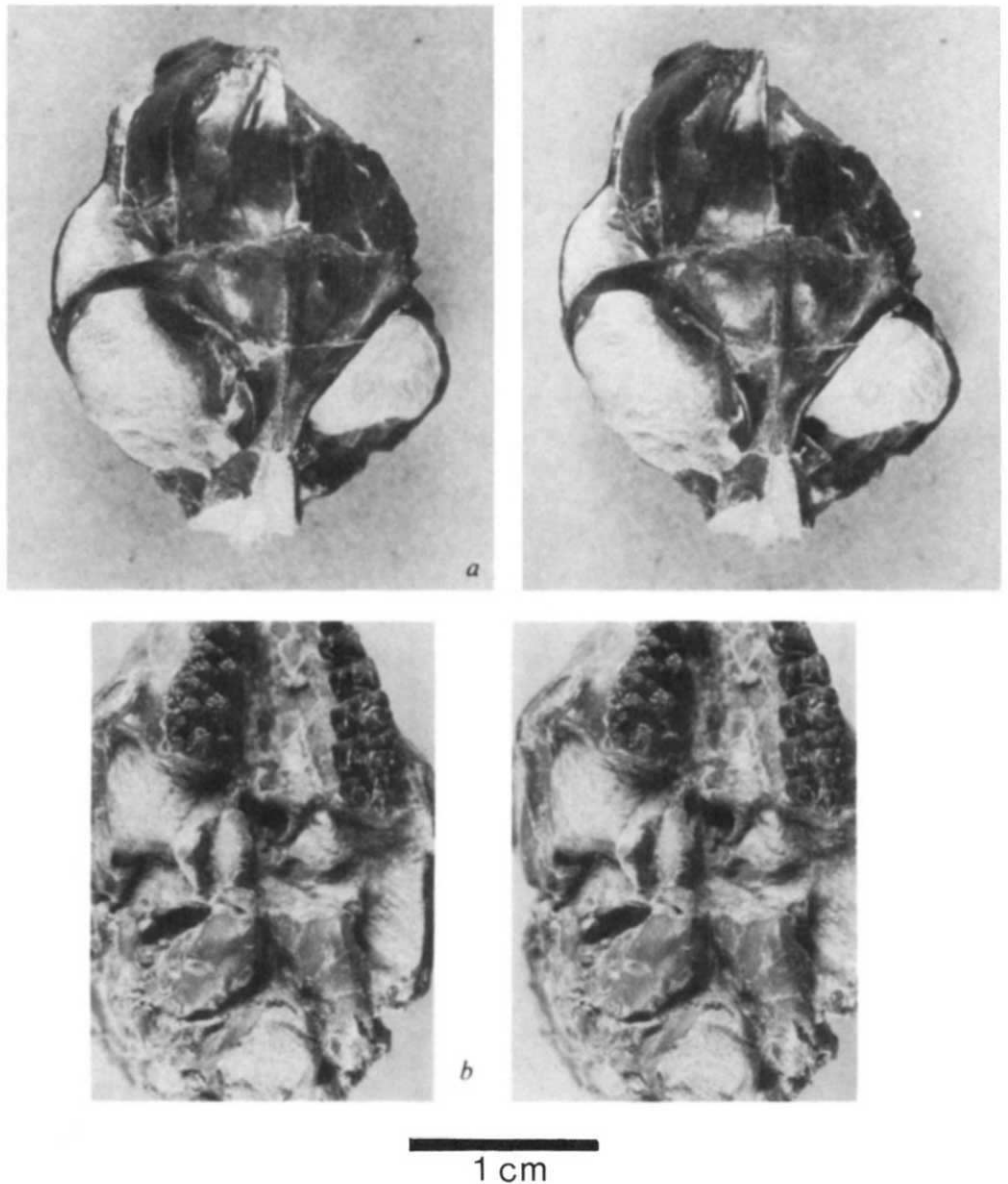


FIG. 2 Reconstructed skull of *Shoshonius cooperi* (based on CM 60494, 60495) in ventral view, showing detail of basicranial region. Abbreviations: ab, annular bridge; ase, anterior intrabullar septum; bos, basioccipital-petrosal suture; eam, external acoustic meatus; fc, facial canal; gf, glenoid fossa; pc, promontory canal; pcf, posterior carotid foramen; pgf, postglenoid foramen; ppg, postglenoid process; prom, promontorium; psep, posterior intrabullar septum; smf, suprameatal foramen; stmf, stylomastoid foramen.

that overlaps the posteromedial bullar wall; (2) a ventrolateral PCF; and (3) presence of an SMF. Other derived features in *Shoshonius* and *Tarsius* are: (1) the lateral pterygoid wings join the anterolateral bullar walls, which also occurs in *Necrolemur* and possibly other Eocene omomyids¹⁰, but not in any other primates; (2) enlarged orbits in proportion to skull length; (3) a reduced snout; and (4) an internal carotid canal that courses vertically within an intrabullar bony septum. The more anterior position of the PCF and the basioccipital flange in *Tarsius* may be the result of strong basicranial flexion of the skull^{4,10,11}, which is lacking in *Shoshonius*.

Considerable palaeontological evidence supports the monophyly of a *Tarsius* + anthropoid + omomyid clade (Haplorhini)^{2-5,7,13,26}; additionally, the *Tarsius*-anthropoid relationship

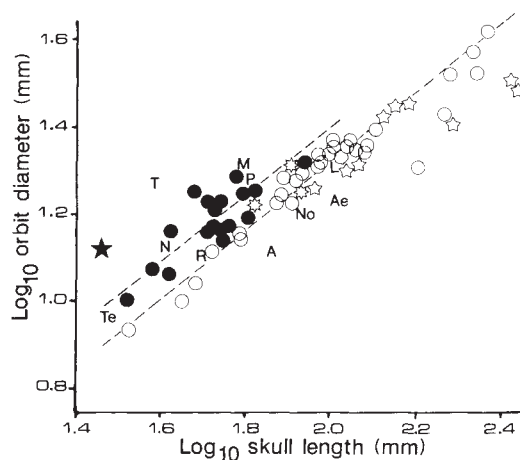


FIG. 3 Logarithmic plot of orbital diameter against skull length in selected living and fossil primates (adapted from Martin³²), including *Shoshonius cooperi* (black star). Other symbols are: A, *Adapis parisiensis*; Ae, *Aegyptopithecus zeuxis*; L, *Leptadapis magnus*; M, *Mioeoticus sp.*; N, *Necrolemur antiquus*; No, *Notharctus tenebrosus*; P, *Pronycticebus gaudreyi*; R, *Rooneyia viejaensis*; T, *Tarsius*; Te, *Tetonius homunculus*; black circles, living nocturnal primates; open circles, living diurnal primates; open starbursts, living crepuscular primates; open stars, subfossil lemurs. Upper line is major axis for living nocturnal primates and lower line is major axis for living diurnal primates.

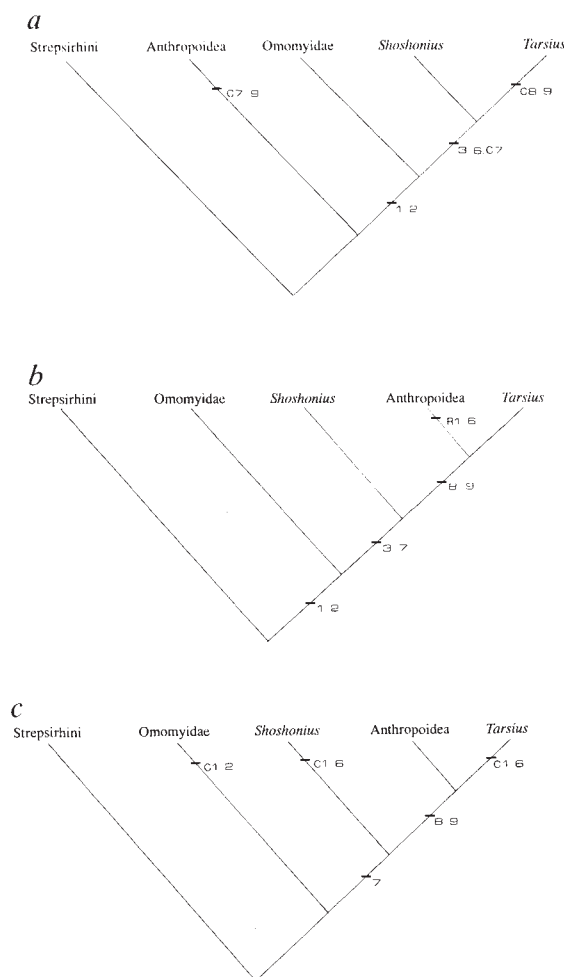


FIG. 4 Alternative hypotheses of phylogenetic relationships among Anthropoidea, *Tarsius*, *Shoshonius* and omomyids known from cranial material (*Necrolemur*, *Rooneyia*, *Tetonius*). a, Most parsimonious PAUP tree: 12 steps, 9 characters, consistency index (CI)=0.750. b, Most parsimonious PAUP tree assuming a *Tarsius*-Anthropoidea clade: 15 steps, CI=0.600, Wagner parsimony allowing character reversals. c, Same tree topology as in b, but not allowing character reversals (Camin-Sokal parsimony): 17 steps, CI=0.529. A character or group of characters preceded by 'C' are inferred convergence(s); those preceded by 'R' are inferred reversal(s). Derived characters are: 1, lateral pterygoid wings overlap anteromedial bullar wall (probably does not occur in *Rooneyia*); 2, presence of SMF (not reported for *Rooneyia* or *Tetonius*); 3, PCF ventrolateral; 4, reduced snout; 5, basioccipital flange overlaps medial bullar wall; 6, enlarged orbits relative to skull length; 7, internal carotid canal enclosed within intrabullar bony septum; 8, loss of ectotympanic annular bridge; 9, partial postorbital closure.

among living primates is bolstered by soft anatomy and molecular evidence^{1,27-32}. The morphology of the new skulls reported here strongly suggests that, within the Haplorhini, *Tarsius* is more closely related to *Shoshonius* than to anthropoids or other Eocene omomyids for which skulls are known (Fig. 4a). This conclusion has three important ramifications for primate evolution.

First, the new evidence reinforces the view¹¹ that three cranial features (partial postorbital closure, possession of an intrabullar bony septum containing the internal carotid canal, loss of the ectotympanic annular bridge) evolved convergently in *Tarsius* and Anthropoidea. In contrast, an anthropoid-*Tarsius* clade that excludes *Shoshonius* and other omomyids^{2-4,12} is less parsimonious: it requires either six reversals to the primitive condition in anthropoids (Fig. 4b) or massive homoplasy in these characters among *Tarsius*, *Shoshonius* and other omomyids (Fig. 4c).

Second, the sister-group relationship between *Tarsius* and *Shoshonius* inferred from the new fossil evidence implies that the family Omomyidae is paraphyletic. Finally, this relationship requires that the divergence of tarsiers and anthropoids occurred before 50.5 Ma, the latter part of the early Eocene. Accordingly, the fossil record of at least the first 15 Ma of anthropoid evolution must be lacking, because the oldest known anthropoid fossils are late Eocene¹⁵⁻¹⁷. □

Received 28 August; accepted 16 October 1990.

- Pocock, R. I. *Proc. zool. Soc. Lond.* **1918**, 19-53 (1918).
- Cartmill, M. & Kay, R. F. in *Recent Advances in Primatology* Vol. 3, *Evolution* (eds Chivers, D. J. & Joysey, K. A.) 205-214 (Academic, London, 1978).
- Cartmill, M., MacPhee, R. D. E. & Simons, E. L. *Am. J. phys. Anthropol.* **56**, 3-21 (1981).
- MacPhee, R. D. E. & Cartmill, M. in *Comparative Primate Biology* Vol. 1, *Systematics, Evolution, and Anatomy* (eds Swindler, D. R. & Erwin, J.) 219-275 (Liss, New York, 1986).
- Rosenberger, A. L. & Szalay, F. S. in *Evolutionary Biology of the New World Monkeys and Continental Drift* (eds Ciochon, R. L. & Chiarelli, A. B.) 139-157 (Plenum, New York, 1980).
- Rasmussen, D. T. *J. hum. Evol.* **15**, 1-12 (1986).
- Szalay, F. S. *Bull. Am. Mus. nat. Hist.* **156**, 157-450 (1976).
- Gingerich, P. D. *J. hum. Evol.* **10**, 345-374 (1981).
- Simons, E. L. *Bull. Br. Mus. (nat. Hist.) Geol.* **5**, 43-69 (1961).
- Rosenberger, A. L. *Folia Primatol.* **45**, 179-194 (1985).
- Simons, E. L. & Rasmussen, D. T. *Am. J. phys. Anthropol.* **79**, 1-23 (1989).
- Cartmill, M. in *Evolutionary Biology of the New World Monkeys and Continental Drift* (eds Ciochon, R. L. & Chiarelli, A. B.) 243-274 (Plenum, New York, 1980).
- Schmid, P. in *Primate Evolutionary Biology* (eds Chiarelli, A. B. & Corruccini, R. S.) 6-13 (Springer, Berlin, 1981).
- Schwartz, J. H. & Tattersall, I. *J. hum. Evol.* **16**, 23-40 (1987).
- de Bonis, L., Jaeger, J.-J., Coiffait, B. & Coiffait, P.-E. *C. r. hebd. Seanc. Acad. Sci., Paris (Ser. II)* **306**, 929-934 (1988).
- Simons, E. L. *Proc. natn. Acad. Sci. U.S.A.* **86**, 9956-9960 (1989).
- Simons, E. L. *Science* **247**, 1567-1569 (1990).
- Stucky, R. K., Krishtalka, L. & Redline, A. D. *Geol. Soc. Am. spec. Pap.* **243**, 169-186 (1990).
- Krishtalka, L. et al. in *Cenozoic Mammals of North America: Geochronology and Biostratigraphy* (ed. Woodburne, M. O.) 77-117 (Univ. California Press, Berkeley, 1987).
- MacPhee, R. D. E. *Contr. Primatol.* **18**, 1-282 (1981).
- Szalay, F. S. & Wilson, J. A. *Folia Primatol.* **25**, 288-293 (1976).
- Conroy, G. C. *Am. J. phys. Anthropol.* **52**, 443-451 (1980).
- Simons, E. L. & Bown, T. M. *Nature* **313**, 475-477 (1985).
- Fleagle, J. G. & Kay, R. F. *J. hum. Evol.* **16**, 483-532 (1987).
- Ginsburg, L. & Mein, P. C. *C. r. hebd. Seanc. Acad. Sci., Paris (Ser. II)* **304**, 1213-1215 (1987).

26. Gebo, D. L. *J. hum. Evol.* **15**, 421–430 (1986).
 27. Luckett, W. P. in *Phylogeny of the Primates: A Multidisciplinary Approach* (eds Luckett, W. P. & Szalay, F. S.) 157–182 (Plenum, New York, 1975).
 28. Pollock, J. I. & Mullin, R. J. *Am. J. phys. Anthropol.* **73**, 65–70 (1987).
 29. de Jong, W. W. & Goodman, M. *J. hum. Evol.* **17**, 575–582 (1988).
 30. Koop, B. F. et al. *J. biol. Chem.* **264**, 68–79 (1989).
 31. Czelusniak, J. et al. in *Current Mammalogy, Volume 2* (ed. Genoways, H. H.) 545–572 (Plenum, New York, 1990).
 32. Martin, R. D. *Primate Origins and Evolution: A Phylogenetic Reconstruction* (Princeton Univ. Press, New Jersey, 1990).

ACKNOWLEDGEMENTS. We thank A. Tabrum for fossil preparation; M. McNaughton for photography; A. Redline for illustrations and technical assistance; R. W. Thorington for the loan of *Tarsius* material; M. Cartmill, R. D. E. MacPhee and R. D. Martin for discussion; and Z. Fross, S. Rose, M. Swarts, D. Swarts, A. Tabrum and A. Redline for fieldwork. This research was supported by the National Science Foundation and the M. Graham Netting Research Fund, Carnegie Museum of Natural History.

Novel form of long-term potentiation produced by a K^+ channel blocker in the hippocampus

L. Aniksztejn & Y. Ben-Ari

INSERM U29, Hôpital de Port-Royal, 123 Boulevard de Port-Royal, 75014 Paris, France

LONG-term potentiation (LTP) of synaptic transmission in the hippocampus is a widely studied model of memory processes¹. In the CA1 region, LTP is triggered by the entry of Ca^{2+} through N-methyl-D-aspartate (NMDA) receptor channels and maintained by the activation of Ca^{2+} -sensitive intracellular messengers^{2,3}. We now report that in CA1, a transient block by tetraethylammonium of I_C , I_M and the delayed rectifier (I_K) produces a Ca^{2+} -dependent NMDA-independent form of LTP. Our results suggest that this new form of LTP (referred as to LTP_K) is induced by a transient enhanced release of glutamate which generates a depolarization by way of the non-NMDA receptors and the consequent activation of voltage-dependent Ca^{2+} channels.

Bath application of tetraethylammonium (TEA) for 10 min at a concentration which blocks I_C , I_M and the delayed rectifier I_K (refs 4, 5 and see also ref. 6 for review), produced an LTP of synaptic transmission. The enhancement of the slope of the field excitatory postsynaptic potential (e.p.s.p.) produced by 25 mM of TEA was $51 \pm 10\%$ ($n = 23/30$), 2 h after wash (Fig. 1a). The enhanced e.p.s.p. persisted for up to 4 h (longest duration of recording performed, $38 \pm 12\%$, $n = 4/4$). There was a significant shift to the left of the input-output curve (Fig. 1b), indicating that after TEA, a given afferent volley will produce a larger postsynaptic response⁷. The threshold concentration of TEA to produce LTP_K was 15 mM ($n = 3$).

Application of TEA at lower concentrations (5 mM), which blocked only I_C and I_M (refs 5, 8 and 9), produced only a transient (5–15 min) increase of the field e.p.s.p. ($n = 3$). Application of other K^+ channel blockers such as 4-aminopyridine (4-AP) (1–2 mM) to block I_A and I_D (ref. 10) ($n = 6$) or caesium (2–3 mM) to block the inward rectifier I_Q (ref. 11) ($n = 6$) produced a decremental potentiation which lasted 45–75 min after washout of the drugs (Fig. 1c–f). With the highest concentration of 4-AP tested (5 mM), full recovery occurred up to 75–90 min after wash ($n = 3$, not shown). In addition, application of noradrenaline (10 μ M) which blocks the slow after-hyperpolarization (I_{AHP})¹², or glibenclamide (5 μ M) which blocks ATP-sensitive K^+ channels (I_{KATP})¹³ had no significant effect on the e.p.s.p. ($n = 3$ respectively, not shown).

To compare the type of excitatory amino-acid receptors involved in LTP_K and electrical LTP, we have tested the effects of selective NMDA and α -amino-3-hydroxy-5-methyl-4-isoxazole propionate (AMPA) receptor antagonists. As shown in Fig. 2a, the selective NMDA receptor antagonist, D-2-amino-

5-phosphonovalerate (APV, 50 or 100 μ M) fully blocked the electrical LTP² but did not prevent LTP_K ($54 \pm 11\%$, 2 h after wash, 5/5). In contrast, the selective AMPA antagonist, 6-cyano-7-nitroquinoxaline-3,3-dione (CNQX)¹⁴, (10 μ M), which blocks excitatory synaptic transmission^{15,16}, prevented LTP_K (Fig. 2b), but not electrical LTP ($n = 6$) (see also refs 17 and 18). Eighty minutes after washout of TEA, a second application of TEA produced an LTP_K in both inputs. Therefore, in contrast to electrical LTP, LTP_K is induced by an enhanced release of glutamate acting on non-NMDA receptors.

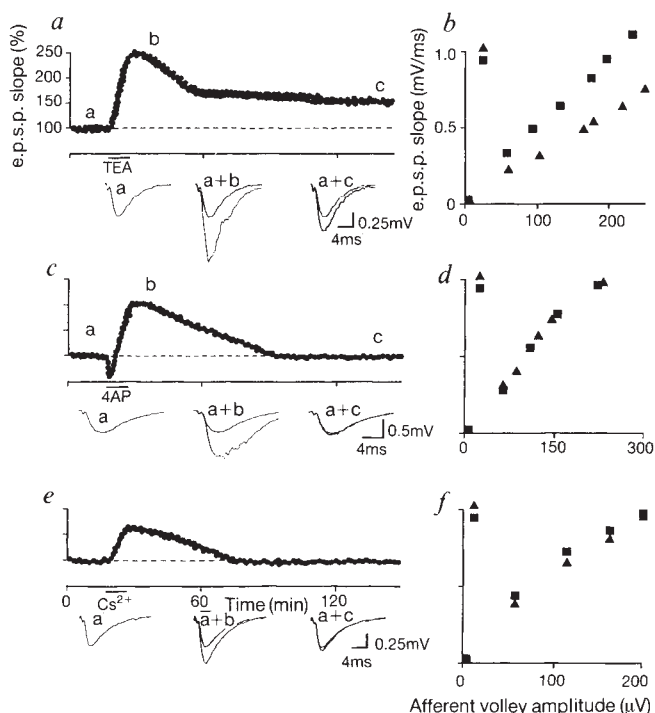
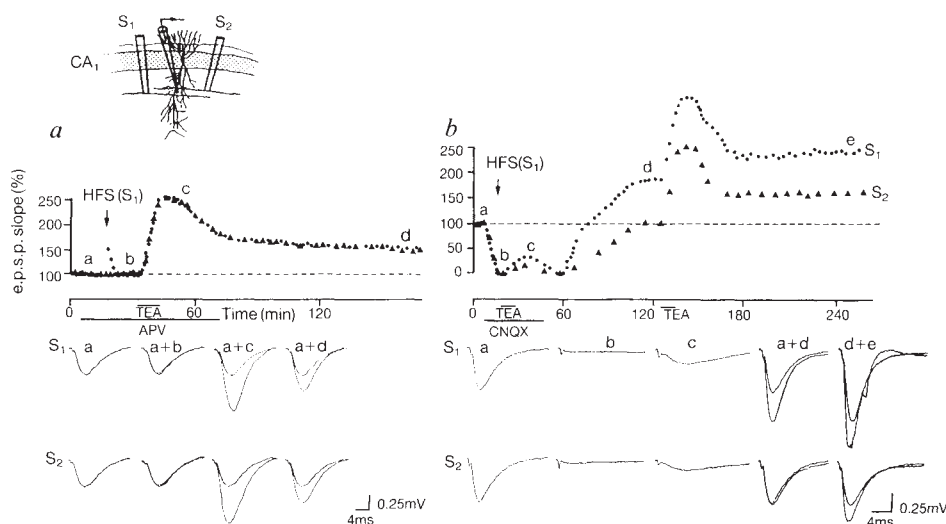


FIG. 1 Effects of TEA, 4-AP, and Cs²⁺ on synaptic transmission. Each graph represents an experiment in which the slope of the field e.p.s.p. has been measured in mV ms⁻¹ and expressed as a percentage against time. Below each graph, representative e.p.s.p.s are shown recorded at a time marked by a letter. On the right side, the slope of the field e.p.s.p.s is plotted against the amplitude of the afferent volley before and 60 min after drug application. Drugs were applied for 10 min. a–b, TEA (25 mM) produced a persistent potentiation of the field e.p.s.p. and a shift of the inward/outward curve. c–f: 4-AP (2 mM, c–d) and Cs²⁺ (3 mM, e–f) produced only a transient potentiation of the field e.p.s.p., full recovery to control response occurred within 60–75 min. b, d, f, Δ , control; $\blacksquare$, 60 min wash.

METHODS. Conventional hippocampal slices were prepared from adult male Wistar rats and maintained *in vitro* as previously described³⁰. The CA1 region was isolated from the CA3 by a knife cut to prevent a spread of bursting activity from the latter region. The medium was superfused at 2.5–3 ml min⁻¹ at 33 °C and contained (in mM): NaCl (126); KCl (3.5); CaCl₂ (2); MgCl₂ (1.3); NaH₂PO₄ (1.2); NaHCO₃ (25); glucose (10); pH 7.4 when equilibrated with 95% O₂ and 5% CO₂. Tetraethylammonium (Sigma), 4-aminopyridine (Sigma), noradrenaline (Sigma), CsCl₂, glibenclamide (gifts from Dr M. Ashford), flunarizine dihydrochloride (Sigma), D-2-amino-5-phosphonovalerate (APV) (Sigma), 6-cyano-7-nitroquinoxaline-2,3-dione (CNQX) (Tocris) were applied in the perfusing medium. Bicuculline (10 μ M) was added to the medium in both APV and CNQX experiments, to facilitate the induction of the electrical LTP³¹. In these conditions the concentration of divalent cation was increased (CaCl₂/MgCl₂, 4 mM) in order to decrease cell excitability. Furthermore, in the CNQX experiments, glycine (1 μ M) was also added to the medium to prevent a possible block of the glycine allosteric site of NMDA receptors³². Field excitatory postsynaptic potentials (e.p.s.p.s) were evoked by a constant stimulation of the Schaffer collaterals, at 0.05 Hz, 30 μ s duration, with a bipolar electrode placed in the stratum radiatum and recorded in the stratum radiatum of CA1 with an electrode containing NaCl (2 M). The intensity of stimulation was adjusted to evoke field e.p.s.p.s only but not population spikes. Intracellular recordings were performed with electrodes (50–70 M Ω) containing KCl (3 M). The microelectrode was connected to a d.c. amplifier (Axoclamp) with a bridge circuit for current injection.

FIG. 2 CNQX but not APV prevents the induction of TEA-induced LTP. Electrodes were positioned as shown in the diagram; two pathways were alternatively stimulated. APV (50–100 μM) or CNQX (10 μM) were applied respectively 20 min and 10 min before TEA and washed out 30 min after the TEA. Electrical LTP was elicited by two tetani at 100 Hz for 1 s separated by 20 s, and was given at twice the control stimulation strength. The trains (HFS) were applied to one input (S1) just before TEA application, the second input (S2) serving as control. *a*, APV (50 μM) which had no effect on the e.p.s.p., blocked the induction of the electrical LTP(S1), but did not prevent LTP produced by TEA (S1 and S2). In contrast, CNQX which blocked the e.p.s.p., prevented the LTP produced by TEA in S2 ($4 \pm 1\%$, $n=5$) but not that induced by a train of high-frequency stimulation in S1 ($63 \pm 20\%$, $n=6$); a second exposure to TEA, 80 min after washing the first application of TEA, produced a persistent potentiation of



the untetanized pathway ($58 \pm 19\%$, $n=5$) and increased the potentiation of the tetanized pathway ($110 \pm 33\%$, $n=5$).

Intracellular recordings showed that TEA (25 mM) produced a small brief depolarization (6 ± 4 mV, $n=9$) and increased neuronal activity associated with the presence of bursts of Ca^{2+} spikes evoked by the stimulation of the Schaffer collaterals (Fig. 3*a, b*). A depolarizing pulse which generated a Na^+ spike followed by an after-hyperpolarization due to the activation of I_C (refs 8 and 9) (arrow), induced in the presence of TEA a burst of Ca^{2+} spikes similar to that evoked synaptically (Fig. 3*c*). This effect was transient however, as 20–30 min after removal of TEA the same depolarizing pulse evoked a Na^+ spike with duration, amplitude and after-hyperpolarization similar to that obtained before TEA. This suggests that the effects of TEA on cellular excitability have been fully reversed by washing. One hour after washing TEA, the intracellular potentiation of the e.p.s.p. was $60 \pm 20\%$ in five cells; in the remaining four cells tested, the e.p.s.p. was clearly potentiated as the stimulation evoked a spike. This potentiation was not associated with any major alteration of passive properties of the cell, that is, no change in spike threshold and in input resistance (51 ± 12 M Ω before TEA and 53 ± 13 M Ω 30 min after; Fig. 3*d*).

To address the possibility that the sustained calcium activity observed in presence of TEA is responsible for the induction of LTP_K, we tested the effect of the calcium channel blocker flunarizine¹⁹ on this potentiation. Flunarizine (20–30 μM) was applied for 15 min before and during the TEA application. Flunarizine had no significant effect on the synaptic transmission or on the Na^+ spike evoked by a depolarizing pulse applied to the cell (Fig. 4). In contrast, under TEA, flunarizine prevented the burst of Ca^{2+} spikes evoked by a depolarizing pulse or by electrical stimulation of the Schaffer collateral (Fig. 4*a, b*).

Flunarizine also strongly reduced or fully prevented LTP_K (Fig. 4*b, c*). Thus, 30–45 min after wash, the intracellular e.p.s.p. ($n=3/3$) and the extracellular field e.p.s.p. had returned to control values ($8.5 \pm 4\%$, Student's *t*-test $P > 0.1$; $n=6/7$).

In common with electrical LTP^{1–3} or the LTP produced by the bee venom peptide (MCD)²⁰, LTP_K reflects a genuine persistent potentiation of synaptic transmission. the dose dependence of the action of TEA as well as the observation that other K^+ channel blockers failed to produce LTP suggest that the blockade of the delayed rectifier (I_K) by TEA is important in initiating LTP_K. There are at present no more selective blockers of I_K , however, and therefore we cannot exclude the possibility that a block of I_C and I_M at the doses of TEA used also

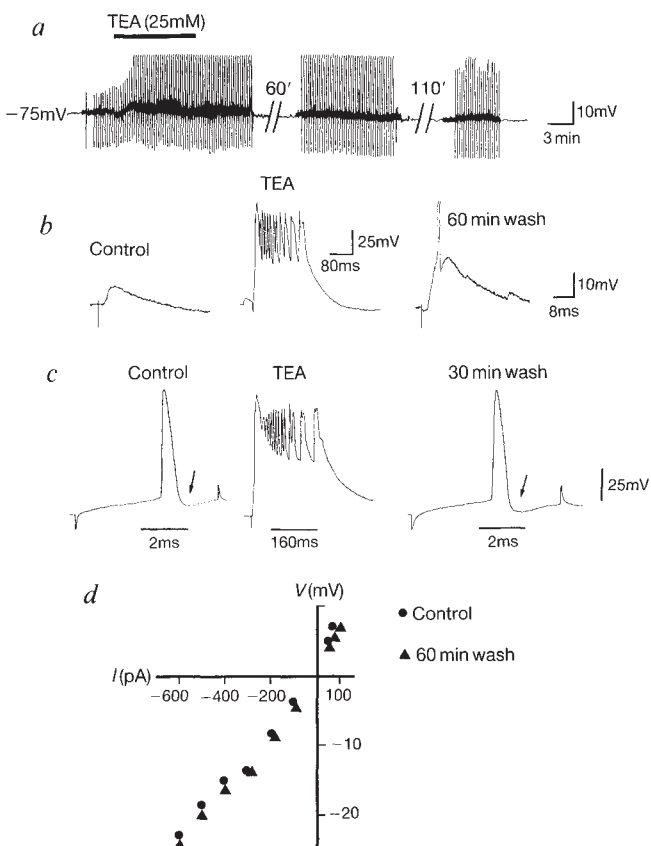
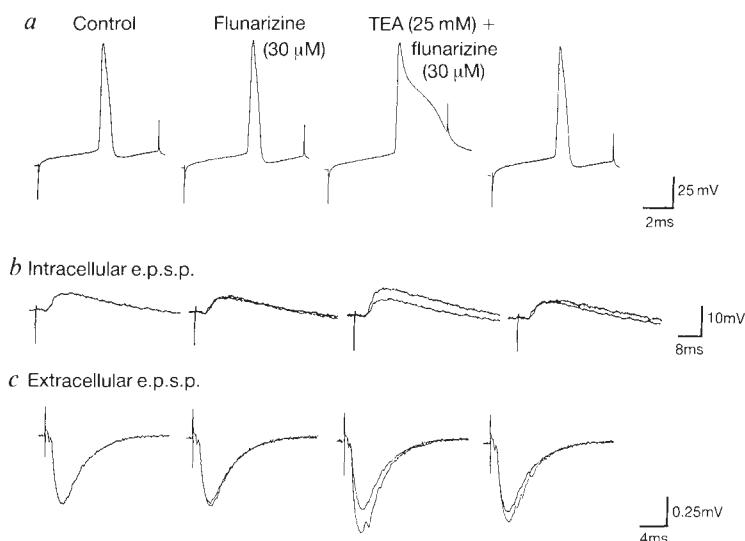


FIG. 3 TEA evoked bursts of Ca^{2+} spikes, but did not produce persistent changes in the passive properties of the cell after wash. *a, b, d* from one cell, *c* from another. *a*, Effect of TEA during continuous recording from a neuron. The upper deflections correspond to the evoked e.p.s.p.s, the lower correspond to electronic pulses produced by a hyperpolarized current pulse (-300 pA, 250 ms duration). Note the brief depolarization and increase in noise during and after TEA. The e.p.s.p. was clearly potentiated 60 min after wash, the electrical stimulation evoked a spike which is truncated. *b*, Representative e.p.s.p. corresponding to *a*. Note the long-lasting burst of Ca^{2+} spikes evoked synaptically under TEA. *c*, Action potential evoked by an intracellular depolarizing pulse (8 ms, 600 pA) before, during and 30 min after TEA. Note the long-lasting burst evoked in the presence of TEA, note also that 30 min after washing TEA, the fast after-hyperpolarization was again observed and the spike duration and shape was similar to that in the control period (arrow). *d*, *V/I* curve corresponding to *a* and constructed before and 60 min after TEA.

FIG. 4 Flunarizine, a calcium channel blocker, prevents the TEA-induced LTP. Flunarizine ($30\ \mu\text{M}$) was dissolved in methanol and applied 15 min before and during TEA exposure. *a*, Flunarizine had no significant effect on the Na^+ spike evoked by a depolarizing pulse (700 pA, 8 ms) but rapidly and completely blocked the long-lasting burst evoked in the presence of TEA. *b*, *c*, Flunarizine did not reduce the intracellularly and extracellularly recorded e.p.s.ps, but blocked LTP_K . Note that flunarizine reduced the large increase in the amplitude of the e.p.s.p. produced by TEA.



contributed to LTP_K . Nevertheless, we suggest that the blockade of these K^+ channels by TEA produces a powerful activation of quisqualate receptors as a result of the large increase in glutamate release. This leads to depolarization at the dendritic level which will facilitate the activation of flunarizine-sensitive voltage-dependent Ca^{2+} channels¹⁹. Under TEA a long-lasting burst of calcium spike was evoked which outlasted the duration of the depolarizing pulse, and so it is possible that the high-threshold slowly activating non-inactivating Ca^{2+} currents are involved. I_K and this I_{Ca} have properties in common, including the threshold potential ($-40\ \text{mV}$ or greater), the kinetics of activation (time constant to peak current is $\leq 180\ \text{ms}$ for I_K , and $100\text{--}300\ \text{ms}$ for I_{Ca}), and no or a slow inactivation^{4,21}. Thus, I_K may efficiently prevent the activation of the high threshold Ca^{2+} currents in physiological conditions. We cannot at present however, exclude an additional presynaptic effect of flunarizine.

It is pertinent to ask why the depolarization produced by TEA is sufficient to stimulate I_{Ca} but not (by removing the voltage-dependent Mg^{2+} block)^{22,23} also to activate the NMDA receptor channel complex. We can exclude a direct block by TEA of NMDA channels²⁴, as in hippocampal slices, NMDA currents are readily generated in the presence of TEA²⁵. Although it cannot be excluded, it is unlikely that the large increase in $[\text{Ca}^{2+}]_i$ produces a desensitization of NMDA receptors²⁶ or a significant reduction in the driving force of I_{NMDA} . Whatever the exact underlying mechanism, our observations suggest that a rise in $[\text{Ca}^{2+}]_i$ not mediated by NMDA receptors can produce an LTP. Recent experiments using Fura-2 imaging suggest that the rise in $[\text{Ca}^{2+}]_i$ produced by a high-frequency train is not entirely NMDA sensitive, raising the possibility of a contribution of voltage-dependent Ca^{2+} channels in the induction of the electrical LTP^{27,28}. The recent study showing the clustering of L-type Ca^{2+} channels in the proximal dendrite of hippocampal pyramidal neurons is in keeping with this possibility²⁹.

Note added in proof: Since the submission of this paper, Grover and Teyler³³ have shown that a high frequency train of electrical stimulation (200 Hz) induced in CA1 a Ca^{2+} -dependent NMDA-independent form of LTP. □

9. Storm, J. F. *J. Physiol., Lond.* **385**, 733–759 (1987).
10. Storm, J. F. *Nature* **336**, 379–381 (1988).
11. Halliwell, J. V. & Adams, P. R. *Brain Res.* **250**, 71–92 (1982).
12. Madison, D. V. & Nicoll, R. A. *J. Physiol., Lond.* **372**, 241–244 (1986).
13. Sturgess, N. C., Ashford, M. L. J., Cook, D. L. & Holes, C. N. *Lancet* **11**, 474–475 (1985).
14. Watkins, J. C., Krogsgaard-Larsen, P. & Honoré, T. *Trends Pharmacol. Sci.* **11**, 25–33 (1990).
15. Blake, J. F., Brown, M. W. & Collingridge, G. L. *Neurosci. Lett.* **89**, 182–186 (1988).
16. Neuman, R. S., Ben-Ari, Y., Gho, M. & Cherubini, E. *Neurosci. Lett.* **92**, 64–68 (1988).
17. Kauer, J. A., Malenka, R. C. & Nicoll, R. A. *Neuron* **1**, 911–917 (1988).
18. Muller, D., Joly, M. & Lynch, G. *Science* **242**, 1694–1697 (1988).
19. Higashi, H., Sugita, S., Matsunari, S. & Nishi, S. *Neuroscience* **34**, 35–47 (1990).
20. Cherubini, E., Ben-Ari, Y., Gho, M., Bidard, J. N. & Lazdunski, M. *Nature* **328**, 70–73 (1987).
21. Brown, T. H. & Griffith, W. H. *J. Physiol., Lond.* **337**, 303–320 (1983).
22. Mayer, M. L., Westbrook, G. L. & Guthrie, P. B. *Nature* **309**, 262–263 (1984).
23. Nowak, L., Bregestowski, P., Ascher, P., Herbert, A. & Prochiantz, A. *Nature* **307**, 462–465 (1984).
24. Wright, J. M. & Nowak, L. M. *Soc. Neurosci. Abstr.* **14**, 1046 (1988).
25. Krnjevic, K., Cherubini, E. & Ben-Ari, Y. *J. Neurophysiol.* **62**, 896–906 (1989).
26. Mayer, M. L. & Westbrook, G. L. *Prog. Neurobiol.* **28**, 197–276 (1987).
27. Regehr, W. G., Connor, J. A. & Tank, D. W. *Nature* **341**, 533–536 (1989).
28. Regehr, W. G. & Tank, D. W. *Nature* **345**, 807–810 (1990).
29. Westenbroek, R. E., Ahljiyanian, M. K. & Catterall, W. A. *Nature* **387**, 281–284 (1990).
30. Gho, M., King, A. E., Ben-Ari, Y. & Cherubini, E. *Brain Res.* **385**, 411–414 (1986).
31. Wigström, H. & Gustafsson, B. *Acta physiol. Scand.* **125**, 159–172 (1985).
32. Kessler, M., Baudry, M. & Lynch, G. *Brain Res.* **489**, 377–382 (1989).
33. Grover, L. M. & Teyler, T. J. *Nature* **347**, 477–479 (1990).

ACKNOWLEDGEMENTS. We thank S. Guidasci for photographic assistance. Financial support from DRET and FIDIA is acknowledged.

Decreased osmotic stability of dystrophin-less muscle cells from the mdx mouse

A. Menke & H. Jockusch

Developmental Biology Unit, University of Bielefeld, POB 8640, W-4800 Bielefeld 1, Germany

HUMAN X-linked Duchenne and Becker muscular dystrophies are due to defects in dystrophin, the product of an exceptionally large gene^{1,2}. Although dystrophin has been characterized as a spectrin-like³ submembranous⁴ cytoskeletal protein, there is no experimental evidence for its function in the structural maintenance of muscle⁵. Current hypotheses attribute necrosis of dystrophin-less fibres *in situ* to mechanical weakening of the outer membrane⁶, to an excessive influx of Ca^{2+} ions^{7,8}, or to a combination of these two mechanisms, possibly mediated by stretch-sensitive ion channels⁹. Using hypo-osmotic shock to determine stress resistance¹⁰ and a mouse model (mdx)^{11,12} for the human disease, we show that functional dystrophin contributes to the stability of both cultured myotubes and isolated mature muscle fibres.

Dystrophic mdx mice were either *mdx/mdx* females or *mdx/y*

Received 1 August; accepted 26 October 1990.

1. Teyler, T. J. & DiScenna, P. A. *Rev. Neurosci.* **10**, 131–161 (1987).
2. Collingridge, G. L. & Bliss, T. V. P. *Trends Neurosci.* **10**, 288–293 (1987).
3. Malenka, R. C., Kauer, J. A., Perkel, D. J. & Nicoll, R. A. *Trends Neurosci.* **12**, 444–450 (1989).
4. Segal, M. & Barker, J. L. *J. Neurophysiol.* **51**, 1409–1433 (1984).
5. Storm, J. F. *J. Physiol., Lond.* **409**, 171–190 (1989).
6. Storm, J. F. *Prog. Brain Res.* **83**, 161–187 (1990).
7. Andersen, P., Sunderberg, S. H., Sveen, O., Swann, J. W. & Wigström, H. *J. Physiol., Lond.* **302**, 463–482 (1980).
8. Lancaster, B. & Nicoll, R. A. *J. Physiol., Lond.* **389**, 187–203 (1987).

males; wild-type (+/+ or +/y) controls were from the same C57BL/10 inbred strain. Intact mature muscle fibres were enzymatically isolated from the interosseus muscle. They were largely devoid of the extracellular matrix but showed normal excitability¹³. There was no significant difference in cell yields and in resting potentials between mdx and control fibres. Using a sheep antibody directed against a fragment of mouse dystrophin (see ref. 14), we showed the peripheral distribution of dystrophin in isolated wild-type interosseus fibres and its absence from corresponding mdx cells.

In response to an abrupt exposure to hypo-osmotic medium (≤ 150 mOsm), the dissociated muscle fibres developed expanding membrane blebs which finally collapsed, while the fibres hyper-contracted to about a third of their original length (Fig. 1). At this stage the fibres were dead by the criterion of trypan-blue uptake. Osmotic cell damage was accompanied by a release of the cytosolic enzyme pyruvate kinase into the medium, indicating excessive membrane leakage.

Based on the proportion of survivors, dissociated interosseus muscle fibres from mdx mice (diameters $32 \pm 7 \mu\text{m}$) were more sensitive to osmotic shock than were wild-type controls ($39 \pm 7 \mu\text{m}$). The difference was found regardless of whether the drop in osmolarity was due to reducing salt or sucrose (Fig. 2a); of whether K^+ was present at 4 or 20 mM; of whether Ca^{2+} (1.5 mM) and Mg^{2+} (1.0 mM) were present, or of whether the temperature was 37 °C or 4 °C (not shown). The stability of fibres from mdx/+ heterozygous females was indistinguishable from that of wild-type controls (Fig. 2a), indicating a functional dominance of dystrophin-expressing nuclei (see refs 14–16). Between postnatal ages 50 and 500 days there was little age dependence of the osmotic stability of interosseus fibres (Fig. 2b). Muscle fibres from mice affected by a different muscle disease, myotonia (phenotype ADR, genotype *adr/adr*; see ref. 17), showed no differences in stability from wild-type (A2G, +/?) control fibres (Fig. 2b). Because the ADR phenotype involves a reduction in chloride conductance and a shift of the fibre type pattern from mostly glycolytic to oxidative¹⁷, this result suggests that osmotic stability is not drastically influenced by these parameters.

As indicated by their centrally located nuclei, nearly all fibres in adult (≥ 80 day) mdx interosseus muscles have arisen by

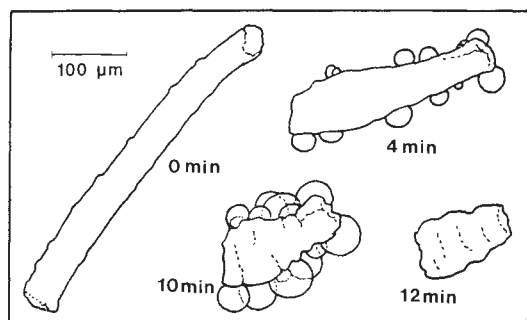


FIG. 1 Response of an isolated wild-type interosseus fibre to hypo-osmotic shock. The times after replacing the 300 mOsm standard medium by 93 mOsm test medium are given. The last (12 min) stage is scored dead. The 'blebbing' reaction was qualitatively similar in mdx fibres but with fewer ($\sim 2/3$) blebs per fibre. Scale bar, 100 μm .

METHODS. Interosseus muscles were dissociated in a rotary shaker (80 r.p.m.) at 37 °C, pH 7.6, with 1.5 mg ml⁻¹ collagenase (type V, from *Clostridium histolyticum*) for 1.5 h and 1.0 mg ml⁻¹ bacterial protease (type IX, from *Bacillus polymyxa*; both from Sigma) for 15 min. Dissociated fibres were allowed to recover for 12 h in growth medium (MEM+20% FCS, 5% CO₂, 37 °C). Hypo-osmotic treatment was by reducing the NaCl of Ca²⁺, Mg²⁺-free phosphate-buffered saline (CMF-PBS; 156 mM Na⁺, 143 mM Cl⁻, 4 mM K⁺, 10 mM phosphate, pH 7.6) at 37 °C.

regeneration (see ref. 18), whereas no central nuclei are found in controls. To distinguish whether previous degeneration and regeneration or the lack of dystrophin are the cause of lowered osmotic stability, interosseus muscles were treated *in situ* with the local anaesthetic bupivacaine, which destroys muscle fibres but not satellite cells, thereby inducing muscle regeneration¹⁹. After 30 to 40 days, very little fibre regeneration (less than 10 per cent of muscle mass) was found in mdx mice, presumably because the pool of satellite cells had been exhausted by previous rounds of necrosis and regeneration. In contrast, wild-type interosseus fibres had regenerated well. After dissociation, their stability was not significantly different from that of non-regenerated fibres (Fig. 2a, b).

Based on microscopic observation and on the release of pyruvate kinase, much lower osmolarities were required to cause membrane damage in cultured myotubes than in mature fibres. Myotubes from mdx mice (diameters $13.6 \pm 3.9 \mu\text{m}$) were more shock-sensitive than control myotubes ($16.1 \pm 3.3 \mu\text{m}$) (Fig. 3). Similarly, no such difference was found between myotubes from

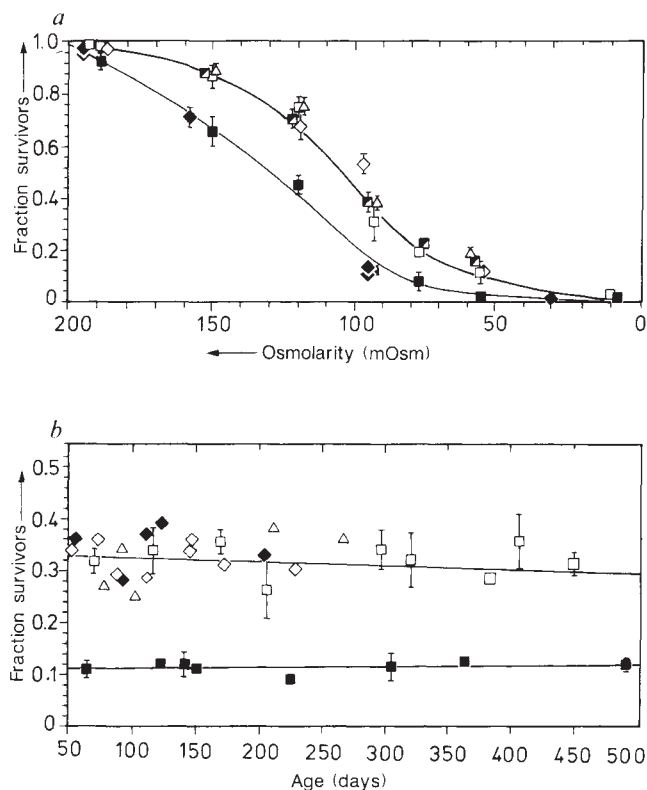


FIG. 2 Resistance to osmotic shock of mature interosseus muscle fibres. **a**, Survival of mdx (■), heterozygous (■), wild-type (□) and regenerated wild-type (△) fibres as a function of osmolarity during 10 min hypo-osmotic treatment; ◇, wild-type, sucrose; ◆, mdx, sucrose. **b**, Dependence of the fraction of surviving fibres after 10 min treatment at 93 mOsm on the age of the animal. In addition to mdx fibres, fibres from myotonic (ADR, *adr/adr*; A2G background) mice were investigated, controls were from +/*adr* or +/+ littermates. Symbols as in **a**, except ◇, wild-type A2G, and ◆, ADR, A2G. **METHODS.** Regenerated wild-type fibres were obtained 40 days after injection of 50 μl 0.75% (v/v) bupivacaine¹⁹ into the interosseus. They had centrally located nuclei like the fibres from adult mdx animals. Hypo-osmotic treatment was in diluted CMF-PBS or in varied sucrose concentrations in the presence of 5 mM Na⁺, 4 mM K⁺ for 10 min at 37 °C and was terminated by the addition of excess 300 mOsm buffer. The fractions of survivors were determined, usually after trypan-blue staining. Mean values ($\pm$ s.d.) are derived from several independent experiments using a total of 14 mdx (5 females and 9 males) and 20 wild-type (9 females and 11 males) mice, as well as single experiments with heterozygous and regenerated wild-type fibres. For each measurement 60 to 250 fibres were counted.

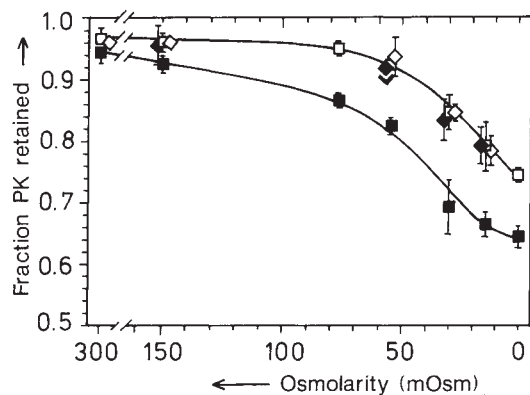


FIG. 3 Osmotic stability of normal and mutant myotubes in culture. Cell damage as a function of osmolarity, as determined by measuring released pyruvate kinase (PK). Retained PK was plotted as an equivalent of cell survival. Model experiments on fibroblast cultures indicated that the contribution to the released PK of contaminating fibroblasts in the primary myotube cultures was negligible. Furthermore, the osmotic stabilities of mdx and wild-type fibroblasts did not significantly differ. □, Wild-type; ■, mdx; C57BL/10; ◆, ADR, A2G; ◇, wild-type, A2G.

METHODS. Myogenic cells were prepared from hind leg muscles of 0–4-day-old mdx, 30–45-day-old ADR and corresponding control mice and cultured according to a protocol modified from ref. 22. When the myotubes showed cross-striations (8–10 days after fusion), the cultures were washed twice with CMF–PBS and treated with diluted CMF–PBS (150–10 mOsm) for 20 min. The release of PK was measured with an assay kit (Boehringer, Mannheim) at 25 °C and related to total releasable PK in the culture, as measured after lysis by freezing in distilled water. Mean values $\pm$ s.d. (mdx, C57BL/10: 8 cultures per point, 6 independent preparations; ADR, A2G: 5 cultures, 3 preparations).

ADR mice and their corresponding A2G wild-type controls, nor were the latter different from the myotubes of the C57BL/10 inbred strain.

The actual mechanism of cell death in our assay system is not known because osmotic stress involves both mechanical tension on and forced ion fluxes across cellular membranes. In mature fibres, the rupture of the outer membrane is probably accompanied by a cascade of intracellular events which lead to hypercontraction. These processes may amplify local damage that occurs anywhere on a surface of several thousand square micrometres of a muscle fibre. Others have attempted to measure membrane stability directly by applying suction through micro-electrodes, thereby sampling a few square micrometres of surface sarcolemma per measurement. No difference from controls was found (by this method) in freshly isolated muscle fibres from the dystrophic dog²⁰, nor in mdx myotubes⁹. Although it is highly suggestive that the difference we measure reflects the presence or absence of dystrophin^{3,4}, other more indirect mechanisms cannot be excluded. For example, the macromolecule conferring stability to the fibres may be normally associated with dystrophin and therefore missing or dislocated in dystrophin-free fibres²¹. Another possibility is that a long-term Ca^{2+} influx into dystrophin-free fibres might, by activating proteases⁸, lead to a degradation of a component involved in fibre stability as measured by hypo-osmotic shock.

We have shown that hypo-osmotic shock applied to mature mouse muscle fibres or to cultured myotubes reveals a hitherto undetected difference between the stability of dystrophin-free mutant cells and their dystrophin-containing counterparts. A rather modest reduction in stability of mdx fibres may well be relevant for the disease. According to ultrastructural observations on the adult mdx mouse, necrosis of a given muscle fibre (or at least a fibre segment in longer fibres) seems to be an all-or-none chance event with a fairly low probability of the

order of a few per cent per week. Equivalent cellular mechanisms may apply to human Duchenne or Becker dystrophies. □

Received 30 August; accepted 25 October 1990.

- Koenig, M. et al. *Cell* **50**, 509–517 (1987).
- Hoffman, E. P., Brown, R. H. & Kunkel, L. M. *Cell* **51**, 919–928 (1987).
- Koenig, M. & Kunkel, L. M. *J. biol. Chem.* **265**, 4560–4566 (1990).
- Cullen, M. J., Walsh, J., Nicholson, L. V. B. & Harris, J. B. *Proc. R. Soc. B* **240**, 197–210 (1990).
- Slater, C. R. *Curr. Opin. Cell Biol.* **1**, 110–114 (1989).
- Karpati, G. & Carpenter, S. *Am. J. Med. Genet.* **25**, 653–658 (1986).
- Duncan, C. J. *Experientia* **45**, 175–177 (1989).
- Turner, P. R., Westwood, T., Regen, C. M. & Steinhardt, R. A. *Nature* **335**, 735–738 (1989).
- Franco, A. & Lansman, J. B. *Nature* **344**, 670–673 (1990).
- Menke, A. & Jockusch, H. *J. neurol. Sci. (sup.)* **98**, 124 (1990).
- Bufffield, G., Siller, W. G., Wight, P. A. L. & Moore, K. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1189–1192 (1984).
- Sicinski, P. et al. *Science* **244**, 1578–1580 (1989).
- Koeltgen, D., Brinkmeier, H. & Jockusch, H. *Muscle & Nerve* (in the press).
- Watkins, S. C., Hoffman, E. P., Slayter, H. S. & Kunkel, L. M. *Muscle & Nerve* **12**, 861–868 (1989).
- Jockusch, H., Friedrich, G. & Zippel, M. *Muscle & Nerve* **13**, 551–555 (1990).
- Partridge, T. A., Morgan, J. E., Gulton, G. R., Hoffman, E. P. & Kunkel, L. M. *Nature* **337**, 176–179 (1989).
- Jockusch, H. in *The Dynamic State of Muscle Fibers* (ed. Pette, D.) 429–443 (de Gruyter, New York, 1990).
- Carnwath, J. W. & Shotton, D. M. *J. neurol. Sci.* **80**, 39–54 (1987).
- Foster, A. H. & Carlson, B. M. *Anesthetics* **58**, 727–736 (1980).
- Cooper, B. J. & Hamill, O. P. *Soc. Neurosci. (Abstr.)* **15**, 412.5 (1989).
- Ervasti, J. M., Ohlendieck, K., Kahl, S. D., Gaver, M. G. & Campbell, K. P. *Nature* **345**, 315–319 (1990).
- Blau, H. M. & Webster, C. *Proc. natn. Acad. Sci. U.S.A.* **71**, 5623–5627 (1981).

ACKNOWLEDGEMENTS. We thank Professor Bücher (Munich) for the mdx stock, R. Boyce and L. M. Kunkel (Harvard Medical School, Boston) for anti-dystrophin antibody, E. Wischmeyer for advice and R. Klocke for typing the manuscript. Supported by the Deutsche Forschungsgemeinschaft.

CD4 expressed on earliest T-lineage precursor cells in the adult murine thymus

Li Wu, Roland Scollay, Mark Egerton*, Martin Pearse*, Gerald J. Spangrude* & Ken Shortman†

The Walter and Eliza Hall Institute of Medical Research, Melbourne, Victoria 3050, Australia

A CONTINUOUS but low input of stem cells or 'prothymocytes' is necessary to maintain T-cell development in the adult thymus¹, but the colonizing cell has not been characterized. Precursors of T cells have been found in the minor CD4⁺8[−] population² of thymocytes, but even the earliest cells of this population already have partially rearranged T-cell antigen receptor (TCR) genes³. We now demonstrate that the thymus contains a minute population of lymphoid cells similar in some but not all respects to bone marrow-derived haemopoietic stem cells. This population has TCR genes in a germline state. It gives a slow but extensive reconstitution of both $\alpha\beta$ and $\gamma\delta$ lineages on transfer into an irradiated thymus, with kinetics indicating that it includes the earliest intrathymic precursor cells so far isolated. Surprisingly, these cells express low surface levels of the mature T-cell marker CD4.

T-lineage precursors in the thymus fall mainly within the heterogeneous group of double negatives (CD4[−]8[−]), which represent 3–5% of adult murine thymocytes^{2,4,5}. We have found evidence³ for a maturation sequence involving subsets of CD4[−]8[−]3[−] T-precursors expressing high levels of heat-stable antigen (HSA). The earlier subsets in this sequence were phenotypically different from the cells in bone marrow capable of reconstituting a thymus^{6–10} and already had partially rearranged TCR γ and β genes. Accordingly, we looked for a still earlier intrathymic precursor.

* Present addresses: Cell Biology & Metabolism Branch, National Institute for Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892, USA (M.E.); Department of Clinical Immunology, St Vincent's Hospital, Fitzroy 3065, Australia (M.P.); Rocky Mountain Laboratories, Hamilton, Montana 59840, USA (G.J.S.).

† To whom correspondence should be addressed.

Thymocyte suspensions were depleted of CD4⁺8⁺ thymocytes and mature T cells (using antibodies against CD8 and CD3 but not CD4) and also of the more mature CD4⁺8⁺ precursors (using antibodies against the interleukin-2 receptor (IL-2R) and CD2 (ref. 11)), and the remaining cells were sorted according to CD4 or HSA expression. Most early precursor activity was found to be associated with cells bearing moderate (but not high) levels of HSA and, surprisingly, with cells bearing moderate (but not high) levels of CD4 (Fig. 1). We shall refer to these cells as low-CD4 precursors. Flow cytometry with another anti-CD4 antibody seeing a different epitope (YTA 3.1.2 (ref. 12)), and polymerase chain reaction analysis for CD4 RNA, confirmed that this was conventional CD4, synthesized by these cells.

To determine in more detail the phenotype of these early T-precursors, we used three-colour immunofluorescent staining and flow cytometry. Visualization of the active subpopulation required extensive prior depletion of the major thymus populations, including the more mature CD4⁺8⁺ precursors (as before). We also depleted non-T lineage cells with antibodies against erythrocytes, granulocytes, macrophages and B cells. A distinct subpopulation was then evident (Fig. 2), representing 15% of the depleted cells and about 0.05% of all thymocytes. This subpopulation expressed intermediate levels of CD4 and HSA, as in Fig. 1, and was also Thy 1⁺ (low), H-2K⁺ (very high) and Pgp-1⁺ (phagocytic glycoprotein-1) (high). Multiple cross correlations in three- or four-colour analysis confirmed that there was a single discrete population with this phenotype. Of the cells in the depleted fractions, only the 15% with the full CD4⁺ Thy 1⁺ HSA⁺ H-2K⁺ Pgp-1⁺ phenotype were active in thymic reconstitution assays, the remainder producing 10⁶-fold fewer progeny at 21 days. This low-CD4 precursor population was also negative for the B-cell marker B220 (ref. 29), in contrast to B-lineage precursors. The optimum combination of markers for sorting, used in all subsequent work, was Thy 1⁺ HSA⁺ H-2K⁺. Antibodies against these markers gave no detectable interference with subsequent biological assays. We will continue to refer to this sorted population as the low-CD4 precursor population.

Analysis of DNA from these purified low-CD4 precursors showed germline configuration for TCR β and γ genes (Fig. 3) and also for IgH genes (data not shown). The lack of detectable

TCR gene rearrangement is further evidence that the low-CD4 precursor population represents the earliest cells so far isolated from the adult thymus.

The rate and extent of thymic reconstitution by the low-CD4 precursors was very different from that given by the CD4⁺8⁺ precursors (Table 1). As shown in Table 1, 7 days after intrathymic transfer the double negative precursors (isolated as CD4⁺8⁺ Pgp-1⁺) had already developed into CD4⁺8⁺ thymocytes, while the progeny of the low-CD4 precursors had only progressed to the CD4⁺8⁺ stage (losing CD4 expression in the process). The low-CD4 precursors took 9 days to produce CD4⁺8⁺ thymocytes; this was still two days faster than bone marrow stem cells¹⁴. The end-product of the developmental sequence, the mature single positives, did not appear in significant quantity until 21 days after thymic reconstitution by the low-CD4 precursors; at this time mature T-cell progeny were also found in spleen and lymph nodes. As well as being much slower, the number of progeny cells produced from a given input was at least 50-fold more from the low-CD4 precursors than from the double-negative precursors (Table 1).

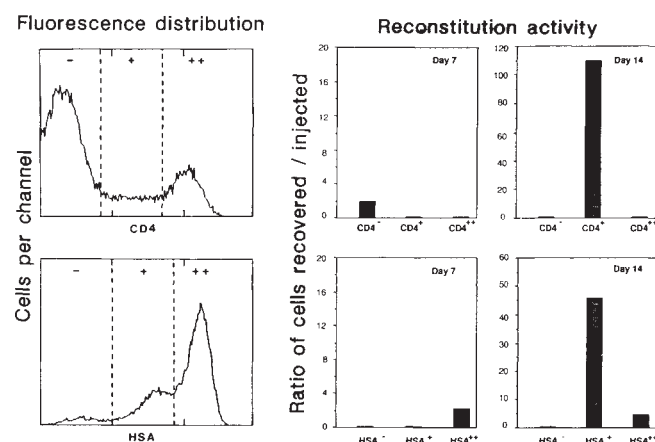


FIG. 1 The levels of CD4 and HSA on early precursor thymocytes. A preparation of C57BL/Ka Thy 1.2 thymocytes depleted of CD8, CD3, CD2, and IL-2R bearing cells was stained with either anti-CD4 (GK1.5) or anti-HSA (M1/69). The distribution of fluorescence and the gates used for subsequent sorting into three fractions are shown. The sorted cells were transferred intrathymically into irradiated C57BL/Ka Thy 1.1 recipients, and the level of Thy 1.2⁺ cells in the thymuses determined 7 or 14 days later. The staining and assay procedures were as in Table 1. Results are the means of 2 experiments, with each point of each experiment involving 4 recipient animals.

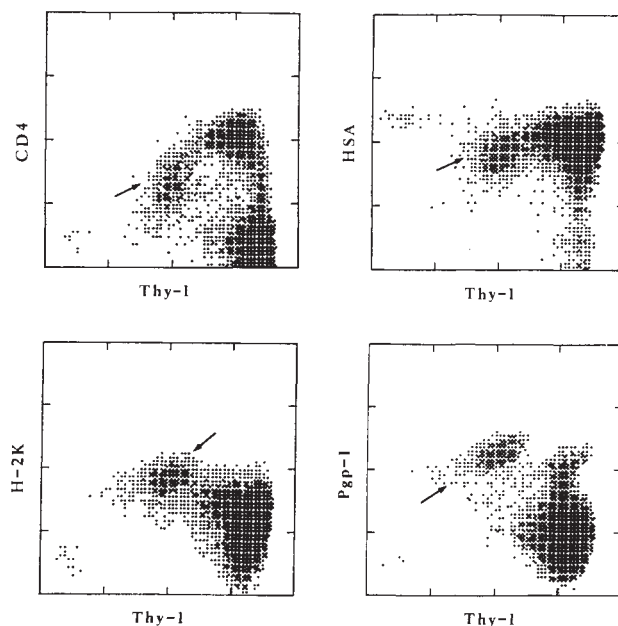


FIG. 2 Surface antigenic phenotype of the low-CD4 precursor thymocyte subpopulation. Thymocyte suspensions were prepared from 16 5-week-old C57BL/Ka Thy 1.2 mice in a HEPES-buffered RPMI 1640 medium containing 10% fetal calf serum, and subjected to an efficient two-step depletion procedure. The first step was treatment with cytotoxic antibodies (30 min, 4°C) then complement (15 min, 37°C), as described elsewhere^{3,4}. The cytotoxic antibodies were: CD8, H02/2 (ref. 19); CD3, 17A2 (ref. 20); IL-2R, 7D4 (ref. 21). Dead cells were then removed by a metrizamide density gradient cut (ref. 22). Residual undepleted cells and cells bearing additional markers were then removed by re-coating the cells with antibodies, followed by depletion with anti-immunoglobulin-coated magnetic beads (Dynabeads, Dynal, Oslo). The monoclonal antibodies used at this step were: CD8, 53.6.7 (ref. 23); CD3, KT3 (ref. 24); IL-2R, PC61 (ref. 25); CD2, RM2-1 (ref. 26) (provided purified by H. Yagita); MAC-1, M1/70 (ref. 27); GR-1, RB6-8C5 (ref. 28); B220, RA3-6B2 (ref. 29); erythroid marker, TER-119 (T. Kina *et al.*, manuscript in preparation; provided by K. Ikuta). The ratio of beads to cells was 8 to 1. The final depleted population was then labelled using various combinations of antibodies against Thy 1, CD4, HSA, H-2K and Pgp-1. Four of the many combinations used are shown. Staining procedures and protocols have been described²². Monoclonal antibodies were as described²², except that anti-H-2K^b was clone B8-24-3 (ref. 30). After staining, at least 25,000 cells were analysed on a FACStar Plus instrument (Becton Dickinson). Dead cells and debris were excluded using low-angle light scatter and propidium iodide staining²². Arrows indicate the position of the one active low-CD4 precursor subpopulation defined by four different criteria. Fluorescence is expressed on a 4-decade logarithmic scale in each case.

TABLE 1 Thymocyte precursor activity

Thymic subset injected	Number injected ($\times 10^{-4}$)	Day of assay	Number recovered ($\times 10^{-4}$)	Ratio recovered/injected	Phenotype of donor cells recovered (%)				
					CD4 ⁺ 8 ⁺	CD4 ⁺ 8 ⁺ Blast	CD4 ⁺ 8 ⁺ Small	CD4 ⁺ 8 ⁺ Mature	CD4 ⁺ 8 ⁺
CD4 ⁺ 8 ⁺ Pgp-1 ⁺ (CD4 ⁺ 8 ⁺ precursors)	50.0	7	473	9	<1	36	60	<1	<1
	50.0	14	257	5	<1	5	59	31	5
CD8 ⁺ 3 ⁺ 2 ⁺ IL-2R ⁺	4.5	7	5	1	83	<5	0	0	0
	2.0	9	9	4	70	11	4	<1	<1
Thy-1 ⁺ HSA ⁺ H-2K ⁺ (low-CD4 precursors)	4.5	11	226	50	8	24	59	<1	<1
	2.0	14	847	423	<1	22	71	<1	<1
	2.0	21	1,010	505	<1	4	79	8	4

Details of the thymus reconstitution assay are given elsewhere⁴. Subpopulations were isolated from 16 5-week-old C57BL/Ka Thy 1.2 mouse thymuses by depletion procedures, followed by sorting for the low CD4 subpopulation, as in Fig. 2. Recipients were 750 rads γ -irradiated 8-week-old C57BL/Ka Thy 1.1 mice; four recipients were set up per time point per experiment. The donor cells were injected intrathymically into one lobe, and this lobe removed for assay at the times stated. The recipient thymus suspensions were stained in three fluorescent colours for Thy 1.2, CD4 and CD8, and donor cells analysed by gating for Thy 1.2⁺ cells⁴. Low-angle light scatter was used to distinguish blasts from small cells, and low-angle light scatter and Thy 1 level used to distinguish true mature thymocytes from the large, Thy 1⁺ immature single positives⁶. For simplicity, immature single positives are not included here. Results are the pooled data from a series of experiments each involving two low-CD4 precursor and one CD4⁺8⁺ precursor time point. Each result is the mean from 2–3 experiments.

These data reflect the development of the major T-lymphocyte lineage expressing the $\alpha\beta$ form of the TCR. But in experiments like those shown in Table 1, 14% of the donor-derived CD4⁺8⁺ thymocytes were $\gamma\delta$ TCR⁺ two weeks after reconstitution. So the low-CD4 population had precursor activity for both the $\alpha\beta$ and the $\gamma\delta$ lineages.

We have not determined the proportion of our thymic population that is active in precursor assays, but their homogeneity by surface phenotype and close similarity to haemopoietic stem cells encourages the view that most are early T-lineage cells. We consider these low-CD4 precursors to represent a newly defined phase of T-cell development in the thymus¹⁵, between the bone marrow-derived stem cells continuously colonizing the

organ and the CD4⁺8⁺ precursors engaged in TCR gene rearrangement. Although they share many surface markers with haemopoietic stem cells, the low-CD4 precursors seem more mature in several respects: many of them are dividing within the thymus (DNA histogram analysis³ showed 14–16% of the population in the S+G₂+M phases of the cell cycle), they progress faster to mature T cells on intrathymic transfer, and they have acquired Sca-2 (stem cell antigen 2; data not shown), a marker absent on purified bone marrow stem cells⁹. On the other hand, they are less mature than the established CD4⁺8⁺ precursors in reconstitution behaviour and TCR gene-rearrangement status. The isolation of these cells from a normal thymus agrees with our previous experiments in which we injected purified bone marrow stem cells into an irradiated thymus. In that case a full 7 days of expansion, corresponding to about 10 cell divisions, occurred before the first recognizable CD4⁺8⁺ subpopulation (HSA⁺ IL-2R⁺) appeared, but during this expansion phase the progeny of the stem cells acquired Sca-2 (ref. 14).

The minute proportion of all thymocytes involved in this early developmental phase, and the fact that the cells express CD4, explains why it has not been recognized previously. It is interesting that low levels of CD4 have been reported¹⁶ on murine bone-marrow multipotential haemopoietic stem cells. This CD4 expression apparently extends to the earliest thymocytes, after which it is lost at the CD4⁺8⁺ precursor stage, then regained at the CD4⁺8⁺ cortical stage. If CD4 is also expressed by the earliest T-precursor cells in the human thymus, its role in the aetiology of acquired immunodeficiency syndrome (AIDS) must be considered¹⁷. In addition, replication of simian immunodeficiency virus in mature T cells is restricted to a subset of CD4⁺ cells bearing the equivalent of Pgp-1 (CD44) (ref. 18). The coincidence of these two markers of HIV-susceptible mature T cells being present on the earliest T-lineage cells in the thymus indicates that they might also be targets for HIV infection. □

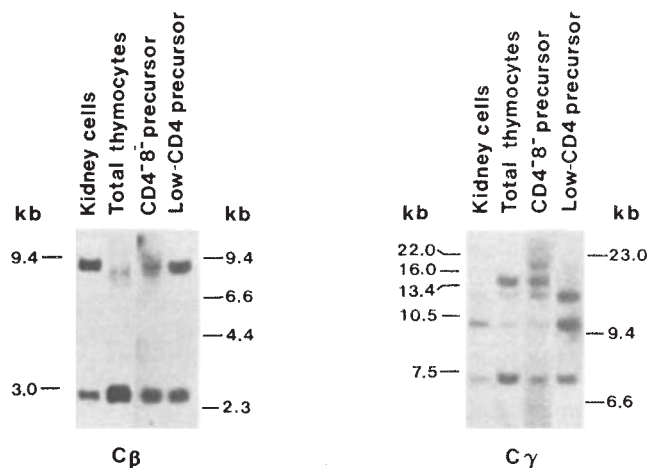


FIG. 3 Status of TCR β and γ -genes. The low-CD4 precursor cells were isolated from 5-week-old C57BL/6 mice, as CD8⁺3⁺2⁺ IL-2R⁺ Thy 1⁺ HSA⁺ H-2K⁺ thymocytes depleted of non-T-lineage cells, according to the procedures used for Fig. 2. Four individual preparations were pooled to provide 5×10^5 cells for each analysis. DNA was prepared for Southern blot analysis as described^{3,13}. The CD4⁺8⁺3⁺ IL-2R⁺ preparation, used as an example of a CD4⁺8⁺ precursor, was prepared as before³. For TCR β -gene rearrangement, DNA was digested with *Hind*III and hybridized with a $C\beta$ probe. Kidney DNA served as a germline control, showing two germline bands of 9.4 kilobases (kb) ($C\beta$ 1) and 3.0 kb ($C\beta$ 2); as only the 9.4-kb band changed on rearrangement, the ratio of intensity of these two bands served as a rough measure of the extent of rearrangement. This ratio was 2.7 for the low-CD4 precursor, compared with a range of 2–3 for various samples of kidney DNA. For TCR γ -gene rearrangement, DNA was digested with *Eco*RI and hybridized with a $C\gamma$ probe. Kidney DNA served as a germline control, showing three germline bands at 13.4 kb, 10.5 kb and 7.5 kb. $C\gamma$ rearrangement was detected by the appearance of discrete additional bands.

Received 6 September; accepted 22 October 1990.

1. Scollay, R., Smith, J. & Stauffer, V. *Immunol. Rev.* **91**, 129–157 (1986).
2. Fowlkes, B. J., Edison, L., Mathieson, B. J. & Chused, T. M. *J. exp. Med.* **162**, 802–822 (1985).
3. Pearce, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 1614–1618 (1989).
4. Scollay, R. *et al. Immunol. Rev.* **104**, 81–120 (1988).
5. Crispe, I. N. *et al. Nature* **329**, 336–339 (1987).
6. Trowbridge, I., Lesley, J., Schulte, R., Hyman, R. & Trotter, J. *Immunogenetics* **15**, 299–312 (1982).
7. Hyman, R., Lesley, J., Schulte, R. & Trotter, J. *Cell Immun.* **101**, 320–327 (1986).
8. Lesley, J., Hyman, R. & Schulte, R. *Cell Immun.* **91**, 397–403 (1985).
9. Spangrude, G. J., Heimfeld, S. & Weissman, I. L. *Science* **241**, 58–62 (1988).
10. Spangrude, G. J. & Scollay, R. *Exp. Hematol.* **18**, 920–926 (1990).
11. Yagita, H. *et al. Eur. J. Immun.* **19**, 2211–2217 (1989).
12. Quin, S., Cobbold, S., Tighe, H., Benjamin, R. & Waldmann, H. *Eur. J. Immun.* **17**, 1159–1165 (1987).
13. Pearce, M. & Wu, L. *Immun. Lett.* **18**, 219–224 (1988).

14. Spangrude, G. J. & Scollay, R. *J. Immun.* (in the press).
15. Shortman, K., Egerton, M., Spangrude, G. J. & Scollay, R. in *Seminars in Immunology* Vol. 2 (ed. Sprent, J.) 3–12 (Saunders, London, 1990).
16. Fredrickson, G. G. & Basch, R. S. *J. exp. Med.* **169**, 1473–1478 (1989).
17. Daigleish, A. G. et al. *Nature* **312**, 763–767 (1984).
18. Gallatin, W. M. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 3301–3305 (1989).
19. Gottleib, P. et al. *Immunogenetics* **10**, 545–555 (1980).
20. Miescher, G. C., Schreyer, M. & MacDonald, H. R. *Immun. Lett.* **23**, 113–118 (1989).
21. Malek, T., Robb, R. & Shevach, E. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5694–5698 (1983).
22. Wilson, A., D'Amico, A., Ewing, T., Scollay, R. & Shortman, K. *J. Immun.* **140**, 1461–1469 (1988).
23. Ledbetter, J. A. & Herzenberg, L. A. *Immunol. Rev.* **47**, 63–90 (1979).
24. Tomonari, K. *Immunogenetics* **28**, 455–458 (1988).
25. Ceredig, R., Lowenthal, J., Nabholz, M. & MacDonald, R. *Nature* **314**, 98–100 (1985).
26. Yagita, H., Nakamura, T., Karasuyama, H. & Okumura, K. *Proc. natn. Acad. Sci. U.S.A.* **86**, 645–649 (1989).
27. Springer, T., Galfré, G., Secher, D. S. & Milstein, C. *Eur. J. Immun.* **9**, 301–306 (1979).
28. Holmes, K. L. et al. *J. Immun.* **137**, 679–688 (1986).
29. Coffman, R. L. *Immunol. Rev.* **69**, 5–23 (1982).
30. Köhler, G., Fischer Lindahl, K. & Heusser, C. in *The Immune System* **2**, 202–208 (Karger, Basel, 1981).

ACKNOWLEDGEMENTS. This work was supported by the NH & MRC Australia, the NIH, and by the C. H. Warman Research Fund, Australia. G.J.S. was a Special Fellow of The Leukemia Society of America. We thank A. D'Amico and D. Vreemec for technical assistance, F. Battye, R. Rossi and S. Unithan for help with cell sorting, K. Kelly for assistance with intrathymic injections, P. Lalor for fluorescent antibodies and S. Cory for the J_{α} probe.

Excess β_2 microglobulin promoting functional peptide association with purified soluble class I MHC molecules

Steven Kozlowski*, Toshiyuki Takeshita†, Wolf-Henning Boehncke‡, Hidemi Takahashi‡§, Lisa F. Boyd*, Ronald N. Germain‡, Jay A. Berzofsky† & David H. Margulies*

*Molecular Biology and ‡Lymphocyte Biology Sections, Laboratory of Immunology, National Institute of Allergy and Infectious Diseases and †Molecular Immunogenetics and Vaccine Research Section, Metabolism Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA

T LYMPHOCYTES expressing $\alpha\beta$ receptors recognize antigenic peptide fragments bound to major histocompatibility complex class I (ref. 1) or class II (ref. 2) molecules present on the surface membranes of other cells. Peptide fragments are present in the two available HLA crystal structures^{3,4} and recent data indicate that peptide is required for the stable folding of the class I heavy chain and maintenance of its association with the class I light chain, β_2 -microglobulin (β_2 m), at physiological temperature^{5–7}. To explain how the exogenous peptide used to create targets for cytotoxic cells bearing CD8 antigen¹ could associate with apparently peptide-filled extracellular class I molecules, we hypothesized that stable binding of exogenous peptide to mature class I molecules reflects either the replacement of previously bound peptide during the well documented β_2 m exchange process⁸ or the loading of 'empty' class I heavy chains dependent on the availability of excess β_2 m. In either case, free β_2 m should enhance peptide/class I binding. Using either isolated soluble class I molecules or living cells, we show here that free purified β_2 m markedly augments the generation of antigenic complexes capable of T-cell stimulation.

To evaluate the role of free β_2 m in the generation of functional class I trimers (that is, peptide/class I/ β_2 m complexes) we developed a cell-free system similar to those previously described^{9,10} for testing specific recognition by T cells. Chimaeric soluble H-2D^d molecules (see Fig. 1) were used to coat the wells of plastic plates. The solid phase H-2D^d was

pulsed with a peptide (p18) corresponding to residues 315–329 of the human immunodeficiency virus type 1 (HIV-1) (IYB) gp160 protein¹¹. The H-2D^d-restricted, p18-specific cell hybridoma B4.2.3, cultured in the same wells, responded with inhibition of its growth to H-2D^d pulsed with p18, but not with the control gp160 peptide T1 (ref. 12) (data not shown). The possibility that free β_2 m played a critical part in the formation *in vitro* of peptide-associated class I molecules recognizable by T cells was then examined by adding affinity-purified human β_2 microglobulin (h β_2 m) during the pulsing with p18. This was performed in the absence of fetal calf serum to avoid the effects of bovine serum proteins, including β_2 m, on the peptide loading process. Figure 1a demonstrates that h β_2 m facilitates p18 association with H-2D^d as assessed by subsequent signalling of the T-cell hybridoma. In fact, no hybridoma response was observed when the peptide was incubated with the class I molecules without h β_2 m. The response of a H-2D^d/p18-specific cytotoxic T lymphocyte clone, RT-1 (ref. 13), is similarly enhanced by h β_2 m (data not shown).

Although the β_2 m that facilitated the ability of H-2D^d to interact functionally with p18 was immunoaffinity-purified, it remained possible that some contaminating material might be the active component of the β_2 m preparation. To test this possibility, the h β_2 m was purified further by size-exclusion chromatography under non-denaturing conditions (Fig. 1b), and the ability of the indicated peak fractions to facilitate the formation of active H-2D^d/peptide complexes was determined (Fig. 1c). Those pooled fractions that promoted the functional interaction of p18 with H-2D^d contain the peak of relative molecular mass 10–12,000 that is the β_2 m monomer (fractions 66–67). Other peaks that are multimers or denatured forms are not active, nor are minor peaks in the included volume of the column that would be expected to contain contaminating peptides.

To determine whether the added h β_2 m binds stably to H-2D^d, as would be predicted from β_2 m exchange, and to examine the role of peptide in any such binding, an enzyme-linked immunosorbent assay (ELISA) was performed (Fig. 2a). Serologically detectable h β_2 m binding to the H-2D^d was half-maximal at a h β_2 m concentration of 2×10^{-9} M in the presence or absence of p18. But, substantially more ($3 \times$) h β_2 m was bound in the presence of peptide than in its absence, suggesting that the association of the H-2D^d heavy chain with peptide is occurring in concert with the binding of h β_2 m. Consistent with this, when the functional response of the T-cell hybridoma is compared with the ELISA data, the dose-response curves correlate well (Fig. 2b). The half-maximal signalling of B4.2.3 occurs at a h β_2 m concentration of $\sim 2 \times 10^{-9}$ M, at which the binding is also half-maximal. Both the h β_2 m effect on peptide pulsing and its binding to H-2D^d occur at concentrations similar to the measured dissociation constant (K_d) of the β_2 m/class I heavy chain interaction, which ranged from 1×10^{-9} – 3×10^{-8} M^{8,14,15}.

Human β_2 m may function in this model system simply by generating additional long-lived 'empty' class I heavy chain/ β_2 m complexes from free heavy chains present in the purified protein preparations that are then able to bind added peptide. If this were so, incubation with β_2 m first, followed by p18, would have the same effect as the simultaneous exposure of the H-2D^d molecule to both components. We tested this possibility by sequentially pulsing the H-2D^d with h β_2 m, followed by p18, or the converse, in addition to our usual co-incubation protocol. Figure 3 summarizes the results of one such experiment, which reveals that simultaneous exposure of the H-2D^d to both p18 and β_2 m is required for the formation of a complex capable of stimulating the antigen-specific H-2D^d-restricted hybridoma. When H-2D^d is simultaneously pulsed with p18 and h β_2 m, even after being washed and incubated without p18 or h β_2 m for the second pulsing interval, it is still able to signal the hybridoma. This result eliminates the possibility that additional incubation time and washes cause the loss of activity of the sequentially pulsed H-2D^d.

§ Present address: Department of Microbiology and Immunology, Nippon Medical School, Tokyo 113, Japan.

FIG. 1 *a*, β_2 -microglobulin dependence of peptide pulsing of H-2D^d assessed by growth inhibition of B4.2.3. (—○—), Immobilized H-2D^d pulsed with p18 in 2% BSA (—●—), immobilized H-2D^d pulsed in 2% BSA with 3 μ g β_2 m (h β_2 m) per well. No peptide, 192,702 c.p.m. $\pm$ 12,911 s.e.m.; no peptide + h β_2 m, 178,269 c.p.m. $\pm$ 18,000 s.e.m. *b*, Human β_2 m (Calbiochem) fractionated by size-exclusion chromatography; shaded areas represent fractions taken for functional assay. *c*, Effectiveness of h β_2 m fractions in augmenting p18 pulsing of H-2D^d-coated plates. Shaded bars, no peptide 18 added; hatched bars, 1 μ M p18 added; assayed by growth inhibition of B4.2.3. Background without hybridoma cells was below 100 c.p.m.

METHODS. The B4.2.3 T-cell hybridoma (anti-H-2D^d plus peptide 18) was generated by the fusion of BW1100, an $\alpha\beta$ -negative mutant of the BW5147 thymoma¹⁹, with lymph node cells of BALB/c mice immunized with p18 in complete Freund's adjuvant²⁰. Soluble H-2D^d protein was affinity purified from serum-free supernatants of L cells transfected with an H-2D^d construct consisting of genomic DNA encoding the H-2D^d $\alpha_1\alpha_2\alpha_3$ domains and the 27 C-terminal amino-acid residues of Q10^b (ref. 21; Kozlowski *et al.*, in preparation). The protein was dialysed against PBS, filter-sterilized, and coated onto Dynatech Immulon plates at the indicated concentration in 50 μ l PBS for 2–2.5 h at 37 °C. The plates were washed twice with 200 μ l PBS, blocked for 15–30 min with 100 μ l of 2% BSA or ovalbumin in PBS and then the peptide in the pulsing medium was added (200 μ l) and the plates incubated overnight at 37 °C and 7.5% CO₂. The following day (after pulsing of at least 15 h) the plates were again washed twice with 200 μ l PBS and the indicated number of B4.2.3 hybridoma cells added per well in DMEM supplemented with 10% FCS, 2 mM glutamine, non-essential amino acids, 50 μ g per ml gentamicin, and 5×10^{-5} M 2-mercaptoethanol. The plates were again incubated overnight at 37 °C and 7.5% CO₂. The next day the cells were pulsed with 1 μ Ci [³H]thymidine (ICN) and collected 4–8 h later for counting the amount of incorporated label (see ref. 22). Results are expressed as c.p.m. averaged from triplicates $\pm$ s.e.m. A similar effect was observed with multiple preparations of H-2D^d, although there was some quantitative variation among different preparations in the extent of h β_2 m-dependent stimulation. *a*, 0.25 μ g per well of H-2D^d coated on Immulon 2 plates. The plate-blocking agent was 2% BSA. Pulsing medium consisted of 2% BSA in PBS with or without 3 μ g per well of human β_2 m (Calbiochem #475823). B4.2.3 was added at 1×10^4 cells per well. *b*, 100 μ g of h β_2 m was fractionated on an HPLC system which consisted of a series of 3 gel filtration columns: TSK 2000 SW followed by 2 TSK 3000 SW columns (7.5 mm i.d. $\times$ 30 cm long). The running buffer was 1 $\times$ PBS, pH 7.2; flow rate was 0.5 ml min⁻¹; 1-min fractions were monitored at 280 nm; *p*-amino-benzoic acid as a marker for the included volume of the column elutes in fractions 73–78. *c*, An Immulon 4 plate was coated with H-2D^d protein at 0.5 μ g per well in 50 μ l PBS for 2 h. The plate was blocked with ovalbumin (Sigma) and pulsed for 24 h with peptide 18 at a final concentration of 1 μ M. Pooled HPLC fractions of h β_2 m were added during pulsing at dilutions of 1:20 to give a final volume of 200 μ l. The unfractionated h β_2 m was added at 0.5 μ g per well. 1×10^4 cells per well of B4.2.3 were cultured with pulsed H-2D^d and their growth inhibition assessed. The experiment was repeated three times using two different fractionations. Results are expressed as the average of triplicates $\pm$ s.e.m.

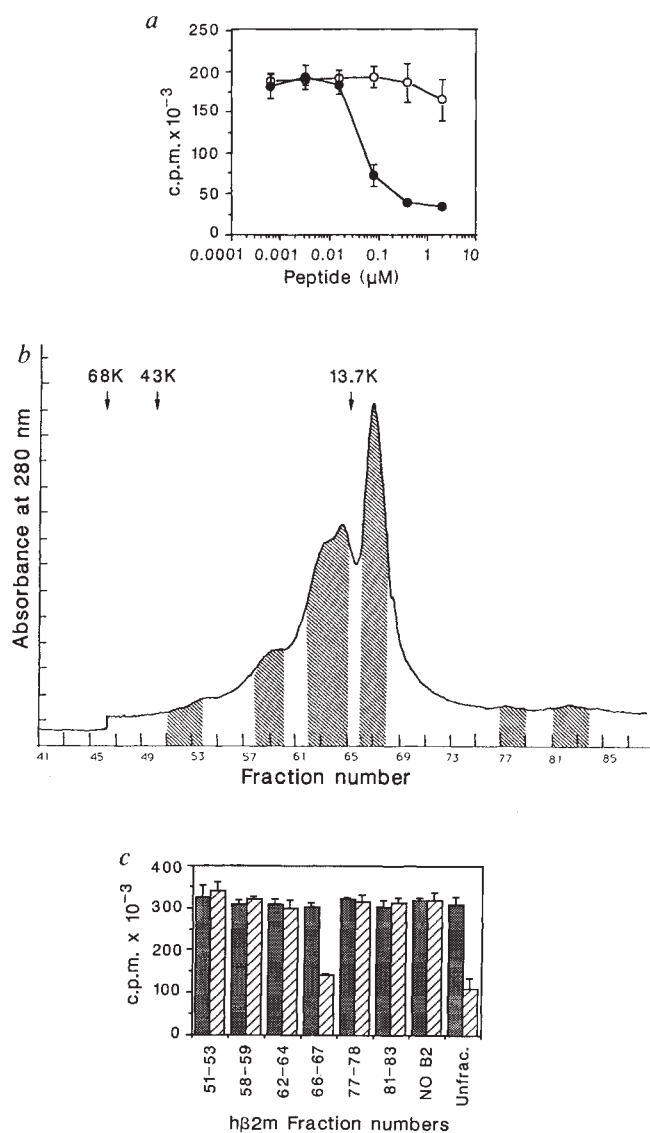
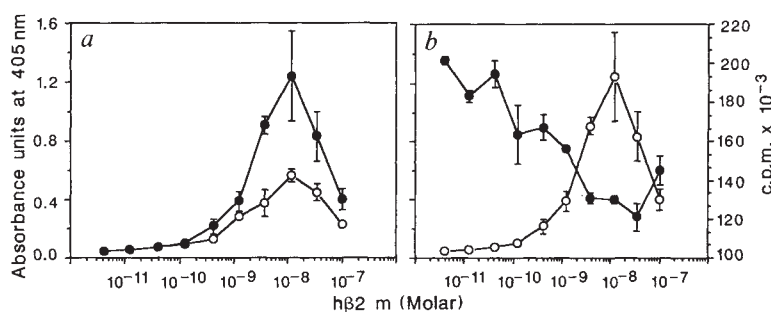


FIG. 2 *a*, ELISA showing binding of h β_2 m to plate-bound H-2D^d. (—○—), Human β_2 m bound without peptide present; (—●—), h β_2 m bound in the presence of 3 μ M p18. Background absorbance was below 0.1 units; h β_2 m added to a blocked plate with or without p18 does not give significant absorbance. *b*, Comparison of h β_2 m binding to immobilized H-2D^d to growth inhibition of B4.2.3. (—○—), Human β_2 m bound in the presence of 3 μ M p18 (405 nm ABS); (—●—), B4.2.3 response (c.p.m.). No h β_2 m = 203,882 c.p.m. $\pm$ 3,116 s.e.m.; no h β_2 m + p18 = 193,790 c.p.m. $\pm$ 8,502 s.e.m.

METHODS. *a*, H-2D^d purified protein was added to an Immulon 4 plate at 0.25 μ g per well in 50 μ l PBS for 2 h at 37 °C. The plate was washed with PBS and blocked with 1.6% BSA, 0.4% ovalbumin in PBS for 30 min. Human β_2 m and p18 (3 μ M) were added as above. After 20 h at 37 °C the plate was washed with PBS Tween 20, 0.02%. For a 2-h incubation, 50 μ l per well of BM-63 ascites (anti-h β_2 m monoclonal, Sigma) diluted 1:400 in 2% BSA, 0.02% Tween-20, PBS was added. The plate was washed again and a rat anti-mouse immunoglobulin conjugated to alkaline phosphatase (Jackson Immunoresearch) was added at a dilution of 1:5,000 in 2% BSA, 0.02% Tween-20, PBS overnight at 4 °C. The assay was developed with Sigma



reagent 104 and absorbance at 405 nm was measured by a Dynatech ELISA reader. Results shown are triplicates $\pm$ s.e.m. Similar results were obtained in three experiments. *b*, The ELISA is as described in *a*. B4.2.3 was added at 10^4 cells per well to a plate treated identically to the ELISA up to the addition of BM-63 ascites. After a 20-h incubation at 37 °C in 7.5% CO₂, the cells were pulsed with [³H]thymidine and collected 4 h later. Results are triplicates $\pm$ s.e.m. Similar results were obtained in three experiments.

The cell-free pulsing system described here consists of affinity-purified class I molecules, affinity-purified and size-fractionated h β 2m, and chromatographically purified peptide in phosphate-buffered saline containing bovine albumin or ovalbumin. The simplicity of the system argues against a requirement for a complicated set of accessory proteins in serum, in the endoplasmic reticulum, or at the cell surface to generate functional class I complexes. The physiological relevance of the data we have obtained in this highly purified system is emphasized by the fact that a similar effect of free h β 2m can be seen when cells bearing surface H-2D^d are incubated with antigenic peptide under serum-free conditions. Figure 4 shows the p18 pulsing of L cells expressing either H-2D^d or H-2K^b with and without h β 2m in serum-free medium. The addition of h β 2m causes a significant augmentation of the H-2D^d-dependent B4.2.3 response, showing that h β 2m-mediated enhancement of peptide/class I association is not limited to the cell-free pulsing system.

The data demonstrate that the functional loading of purified, soluble, mature class I molecules with antigenic peptide is dependent on the simultaneous presence of excess β 2m and peptide in the incubation medium. They also document the enhancement of the binding-exchange of h β 2m onto affinity-purified H-2D^d in the presence of a peptide known to bind H-2D^d. The results suggest a mechanism in which peptide binding is most efficient during the process of β 2m binding-exchange or involves a short-lived bimolecular intermediate. Models of the formation of recognizable class I/ β 2m-peptide complexes that postulate either (1) a long-lived ($t_{1/2}$ of dissociation greater than several minutes at 37°C in the absence of

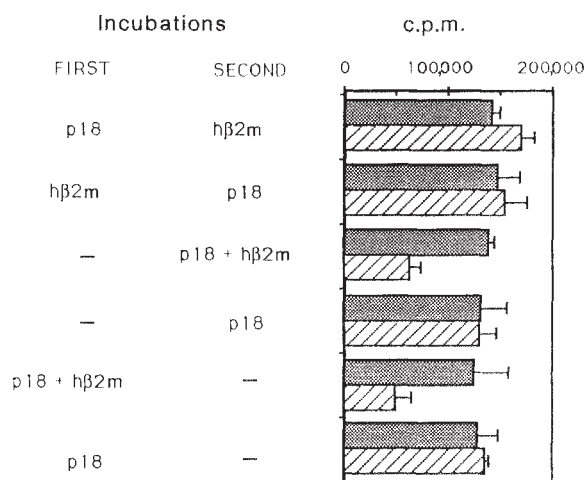


FIG. 3 Sequential pulsing of H-2D^d with p18 and h β 2m. Shaded bars, no peptide added to any groups and h β 2m added as indicated; hatched bars, p18 and h β 2m added as follows: (p18/ β 2), first incubation with 1 μ M p18, second with 0.1 μ g per well h β 2m; (β 2/p18), first incubation with 0.1 μ g per well h β 2m, second with 1 μ M p18; (-p18 + β 2), first incubation with neither, second incubation with both; (-p18), first incubation with neither, second incubation with 1 μ M p18; (p18 + β 2/-), first incubation with both, second incubation with neither; (p18/-), first incubation with 1 μ M p18, second with neither. The background without hybridoma cells was below 100 c.p.m.

METHODS. H-2D^d protein was coated on an Immulon 4 plate at 0.5 μ g per well in 50 μ l PBS for 2.5 h at 37°C. The plate was washed twice, blocked with 2% ovalbumin, and 1 μ M p18, 0.1 μ g h β 2m, both or neither being added in 200 μ l 1% ovalbumin in PBS. After a 23-h incubation at 37°C in 7.5% CO₂ the plate was washed four times (30 s, 200 μ l PBS wash) and the second stage of pulsing was performed as described for 30 h at 37°C in 7.5% CO₂. The plate was washed twice and 1×10^4 cells per well of B4.2.3 added in 200 μ l DMEM with 10% FCS and additives. After 15 h of incubation at 37°C in 7.5% CO₂, the cells were pulsed with 1 μ Ci [³H]thymidine and collected 5 h later. Similar results were obtained in two experiments. Data shown as triplicates $\pm$ s.e.m.

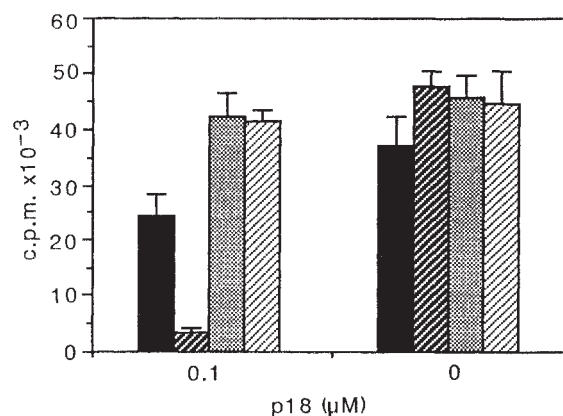


FIG. 4 Human β 2m enhancement of peptide pulsing of H-2D^d-positive L cells. Filled bars, H-2D^d transfected L cells; dark hatched, H-2D^d transfected L cells plus 0.5 μ g per well h β 2m; stippled, H-2K^b transfected L cells; hatched, H-2K^b transfected L cells plus 0.5 μ g per well h β 2m. Background, below 100 c.p.m.

METHODS. DAP-3 cells (thymidine kinase-negative mouse L cells), transfected with a neomycin-resistance gene and either genomic H-2D^d or a H-2K^b cDNA were added to a Costar 96-well plate in 100 μ l DME with 10% FCS at 10^5 cells per well. After an overnight incubation to allow cell adherence, the plate was washed three times with serum-free medium (equal parts Iscove's DMEM and NCTC-109 with 0.4% BSA and 0.1% ovalbumin). Then p18 and h β 2m were added as indicated to give a final volume of 200 μ l. After a 4-h incubation at 37°C in 7.5% CO₂, the plate was again washed three times with serum-free medium. B4.2.3 cells (10^4) were added per well in DMEM, with 10% FCS and additives as in Fig. 1, and more than 15 h later the cells were pulsed with [³H]thymidine. The cells were collected and counted after 5 h. Results are shown as triplicates $\pm$ s.e.m. Similar results were obtained in three experiments.

β 2m) β 2m/heavy chain heterodimer that then binds peptide, or (2) a stable peptide-heavy chain heterodimer that is changed by β 2m into a complex recognizable by the T-cell antigen receptor are inconsistent with our results. Another hypothesis is that β 2m plays the part of both a catalyst for peptide exchange and a ligand for class I heavy chain. An example of this might be formation of a complex intermediate consisting of one heavy chain and two β 2m molecules^{16,17}. The observed reaction order of β 2m exchange-binding, however, is inconsistent with this hypothesis^{8,14}.

Several recent studies have shown the capacity of β 2m (ref. 6), peptide^{6,7}, or low temperature¹⁸ to stabilize the conformation of class I heavy chains in detergent lysates or at the cell surface. Here we have extended this concept of mutual stabilizing interactions among the components of the class I ternary complex⁵ in a cell-free pulsing system using highly purified components.

Note added in proof: K. L. Rock *et al.*^{2,3} have recently published observations similar to ours documenting β_2 microglobulin enhancement of peptide pulsing of H-2K^b-expressing cells. □

Received 28 June; accepted 29 October 1990.

1. Townsend, A. R. M. *et al.* *Cell* **44**, 959-968 (1986).
2. Schwartz, R. H. *et al.* *Rev. Immun.* **3**, 237-261 (1985).
3. Bjorkman, P. J. *et al.* *Nature* **329**, 506-512 (1987).
4. Garrett, T. P., Saper, M. A., Bjorkman, P. J., Strominger, J. L. & Wiley, D. C. *Nature* **342**, 692-696 (1989).
5. Townsend, A. R. M. *et al.* *Nature* **340**, 443-448 (1989).
6. Townsend, A. R. M. *et al.* *Cell* **62**, 285-295 (1990).
7. Schumacher, T. N. M. *et al.* *Cell* **62**, 563-567 (1990).
8. Hyafil, F. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5834-5838 (1979).
9. Kane, K. P., Vitiello, A., Sherman, L. A. & Mescher, M. E. *Nature* **340**, 157-159 (1989).
10. McCluskey, J., Boyd, L. F., Highet, P. F., Inman, J. & Margulies, D. H. *J. Immun.* **141**, 1451-1455 (1988).
11. Takahashi, H. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **85**, 3105-3109 (1988).
12. Cease, K. B. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **84**, 4249-4253 (1987).
13. Takahashi, H. *et al.* *J. exp. Med.* **170**, 2023-2035 (1989).
14. Parker, K. C. & Strominger, J. L. *Biochemistry* **24**, 5543-5550 (1985).
15. Hochman, J. H., Shimizu, Y., DeMars, R. & Edidin, M. *J. Immun.* **140**, 2322-2329 (1988).
16. Ward, P. J. & Sanderson, A. R. *Immunology* **48**, 87-92 (1983).

17. Parham, P. *Hum. Immun.* **8**, 207–215 (1983).
18. Ljunggren, H.-G. *et al. Nature* **346**, 476–480 (1990).
19. White, J. *et al. J. Immun.* **143**, 1822–1825 (1989).
20. Kappler, J. W., Skidmore, B., White, J. & Marrack, P. *J. exp. Med.* **153**, 1198–1214 (1981).
21. Margulies, D. H., Ramsey, A. L., Boyd, L. F. & McCluskey, J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5252–5256 (1986).
22. Ashwell, J. D., Cunningham, R. E., Noguchi, P. D. & Hernandez, D. *J. exp. Med.* **165**, 173–194 (1987).
23. Rock, K. L., Rothstein, L. E., Gamble, S. R. & Benacerraf, B. *Proc. natn. Acad. Sci. U.S.A.* **87**, 7517–7521 (1990).

ACKNOWLEDGEMENTS. We thank M. Corr, R. Ribaudo, R. Hunziker and C. Hoes for their help, S. Starnes for editorial assistance, W. Paul for reviewing the paper, and K. Polin for discussion. W.-H. Boehnke was supported by the Deutsche Forschungsgemeinschaft.

Activation by adrenaline of a low-conductance G protein-dependent K^+ channel in mouse pancreatic B cells

Patrik Rorsman, Krister Bokvist, Carina Ämmälä, Per Arkhammar*, Per-Olof Berggren*, Olof Larsson† & Karin Wählander

Department of Medical Physics, Gothenburg University, Box 33031, S-400 33 Gothenburg, Sweden

* Department of Endocrinology, Karolinska Institute, Box 60500, S-104 01 Stockholm, Sweden

† Permanent address: Department of Pharmacology, Karolinska Institute, Box 60400, S-104 01 Stockholm, Sweden

INSULIN is produced and secreted by the B cells in the endocrine pancreas. *In vivo*, insulin secretion is under the control of a number of metabolic, neural and hormonal substances. It is now clear that stimulation of insulin release by fuel secretagogues, such as glucose, involves the closure of K^+ channels that are sensitive to the intracellular ATP concentration (K_{ATP} channels)¹. This leads to membrane depolarization and the generation of Ca^{2+} -dependent action potentials². The mechanisms whereby hormones and neurotransmitters such as adrenaline, galanin and somatostatin, which are released by intraislet nerve endings and the pancreatic D cells, produce inhibition of insulin secretion are not clear³. Here

we show that adrenaline suppresses B-cell electrical activity (and thus insulin secretion) by a G protein-dependent mechanism, which culminates in the activation of a sulphonylurea-insensitive low-conductance K^+ channel distinct from the K_{ATP} channel.

Addition of adrenaline (5 μ M; Fig. 1a) inhibited glucose-stimulated B-cell electrical activity and repolarized the cell by 20 ± 5 mV ($n=4$). This effect was mimicked by the α_2 -adrenoreceptor agonist clonidine (5 μ M; Fig. 1b) which produced 20 ± 3 mV repolarization ($n=11$). The actions of both adrenaline and clonidine were blocked by the α_2 -adrenoreceptor antagonist yohimbine, consistent with the idea that the effects of adrenaline are mediated by activation of α_2 -adrenoreceptors. Pretreatment of the B cells with pertussis toxin abolished the response to adrenaline (Fig. 1c; $n=4$), indicating the involvement of an inhibitory G protein.

The effects of clonidine on electrical activity is concomitant with an increased resting whole-cell membrane conductance (Fig. 2). In the presence of 20 mM glucose, an input conductance of 0.42 ± 0.11 nS ($n=5$) was observed. Clonidine (5 μ M) increased this by 0.15 ± 0.04 nS ($n=5$; $P<0.05$). Subsequent addition of yohimbine counteracted the effect of clonidine and produced a reduction of 0.16 ± 0.04 nS ($P<0.025$; $n=4$). These effects are small when compared with the total glucose-sensitive K^+ conductance (≈ 5 nS; ref. 6) and the >10 -fold increase in the resting conductance obtained after addition of the K_{ATP} channel activator diazoxide⁷.

Hypoglycaemic sulphonylureas, such as glibenclamide, selectively and effectively block the K_{ATP} channel in pancreatic B cells⁷. Figure 3a shows that 5 μ M glibenclamide fails to restore electrical activity in the presence of clonidine. As K_i for the inhibitory action of glibenclamide on the K_{ATP} channels is in the nano- or subnanomolar range⁸, it seems unlikely that activation of K_{ATP} channels is involved. It is of interest that the repolarizing actions of somatostatin and galanin were also resistant to glibenclamide (not shown). This suggests that in mouse B cells, contrary to what has previously been demonstrated in studies on the insulin-secreting cell line RINm5F^{9,10}, activation of K_{ATP} channels does not account for the hyperglycaemic action of these peptides. The addition of 1 mM tetraethylammonium (TEA), a fairly potent and selective blocker of the Ca^{2+} -activated K^+ channel (K_{Ca} channel) in the B cell¹¹, also fails to restore electrical activity, suggesting that activation (Fig. 3b) of K_{Ca} channels is not involved.

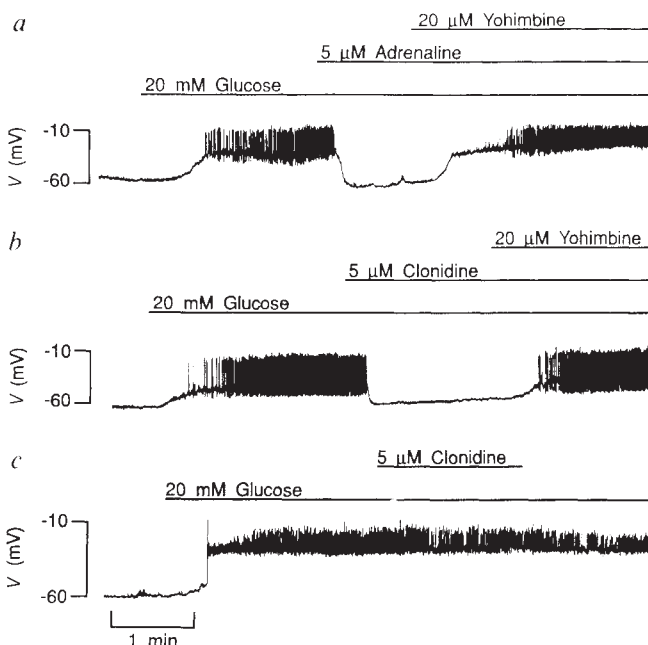


FIG. 1 Effects of adrenergic stimulation on glucose-induced electrical activity recorded from small B-cell clusters. *a*, Effects of 5 μ M DL-adrenaline and 20 μ M yohimbine. *b*, Effects of 5 μ M clonidine and 20 μ M yohimbine. *c*, Pretreatment of cells with pertussis toxin (100 ng ml⁻¹; Sigma) overnight abolishes the repolarizing effect of clonidine.

METHODS. Mouse pancreatic B cells were isolated and maintained in primary tissue culture as described previously⁴. The cells were continuously superfused (bath volume: 0.5 ml; rate: 4 ml min⁻¹) with an extracellular medium containing (in mM) 138 NaCl, 5.6 KCl, 1.2 MgCl₂, 2.6 CaCl₂ and 5 HEPES-NaOH, pH 7.40. Membrane potentials and membrane currents were recorded using the perforated patch configuration of the patch-clamp technique⁵. The pipette-filling solution contained (in mM) 10 KCl, 76 K₂SO₄, 10 NaCl, 1 MgCl₂ and 10 HEPES-KOH, pH 7.35 and 50 μ g ml⁻¹ nystatin (dissolved in dimethylsulphoxide (DMSO); final concentration of DMSO, 0.1%). Yohimbine and glibenclamide were prepared as a 1000 $\times$ concentrated stock solution in DMSO (final concentration of DMSO: 0.1–0.15%). Current and voltage signals were recorded using a List EPC-7 patch-clamp amplifier and stored on magnetic or video-tape pending analysis. Filter settings are given as -3 dB values. All experiments were performed at 29–32°C.

FIG. 2 Parallel recordings of membrane potential and membrane conductance in a single B cell. *a*, Membrane potential recording. Glucose, clonidine, yohimbine and diazoxide were added as indicated by the horizontal bars. During the periods indicated by the asterisks (*), the amplifier was switched from current-clamp mode (membrane potential recording) to the voltage-clamp mode (measurements of membrane currents). The resting conductance observed in the presence of glucose (i) and after addition of clonidine (ii), yohimbine (iii) or diazoxide (iv) was monitored by the application of 20 mV hyper- and depolarizing voltage pulses from a holding potential of -70 mV. The resulting current responses are displayed in *b* and (at greater resolution) in *c*. Current signals were filtered at 200 Hz.

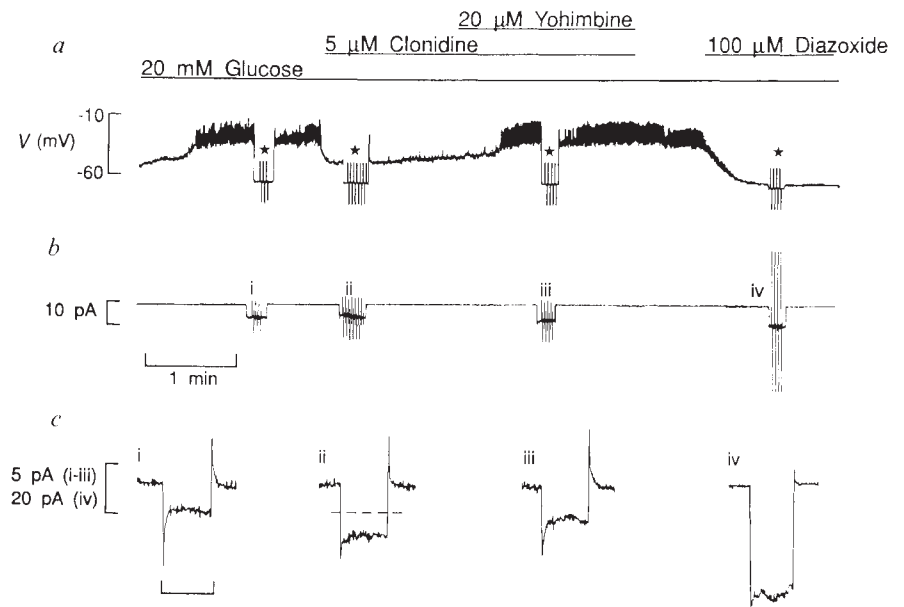


Figure 4a shows a whole-cell voltage-clamp recording from a B cell under basal conditions and after addition of tolbutamide and clonidine. The current observed under basal conditions reflects the summed activity of all the K_{ATP} channels in the entire B cell. Adding tolbutamide ($500 \mu\text{M}$) or glucose (20 mM) blocks this current. Subsequent addition of clonidine (0.1 – $5 \mu\text{M}$) leads to the development of a small outward current, amounting to $11.1 \pm 3.1 \text{ pA}$ ($n = 5$) at $5 \mu\text{M}$. The I – V relationships of the current activated by clonidine at 5.6 mM and 20.6 mM $[K^+]_0$ are shown in Fig. 4b. With normal $[K^+]_0$, a reversal potential of $-74 \pm 4 \text{ mV}$ was observed ($n = 10$). After increase of $[K^+]_0$ to 20.6 mM , the zero current potential was $-42 \pm 5 \text{ mV}$ ($n = 4$). This shift of 32 mV is close to the 34 mV predicted by the Nernst equation for a 3.6-fold increase in $[K^+]_0$, indicating that the channel activated by clonidine is selective for K^+ .

Figure 4c shows expanded segments of the current record under basal conditions (trace i), after addition of tolbutamide (trace ii) and in the presence of clonidine (trace iii). The development of the clonidine-induced current is associated with little excess noise (compare traces ii and iii). The unitary amplitude (i) of the channel activated by clonidine, estimated by stationary fluctuation analysis¹², was $0.03 \pm 0.02 \text{ pA}$ ($n = 5$) from which a linear slope conductance of $\sim 0.6 \text{ pS}$ can be estimated using our value for the reversal potential. The corresponding value of i

under basal conditions (in the absence of tolbutamide and clonidine) was $0.4 \pm 0.2 \text{ pA}$ ($n = 5$). This value is in reasonable agreement with the 0.6 pA obtained during single-channel recordings for the K_{ATP} channel⁴. It seems justifiable to conclude that the amplitude of the clonidine-activated K^+ channel is substantially lower than that of the K_{ATP} channel. Given the value of i and the magnitude of the whole-cell current, a minimum density of the clonidine-activated K^+ channel of 370 copies per cell can be calculated.

It seems unlikely that the effects of adrenaline are attributable to interference with adenylate cyclase as neither application of the adenylate cyclase activator forskolin nor addition of dibutyryl cyclic AMP counteract the repolarizing response to clonidine¹³. In support of a more direct interaction between G proteins and the clonidine-activated K^+ channel, intracellular application of GTP- γ -S leads to the slow (2–3 min) activation of a current with an appearance similar to that observed in response to the agonist (Fig. 4d).

The fact that adrenaline activates K^+ channels distinct from the K_{ATP} channel provides a mechanism whereby the action of neurotransmitters can be dissociated from those of fuel secretagogues. Activation of this K^+ conductance, albeit small, will repolarize the B cell sufficiently to suppress glucose-stimulated insulin secretion. Moreover, any stimulation of neuro-

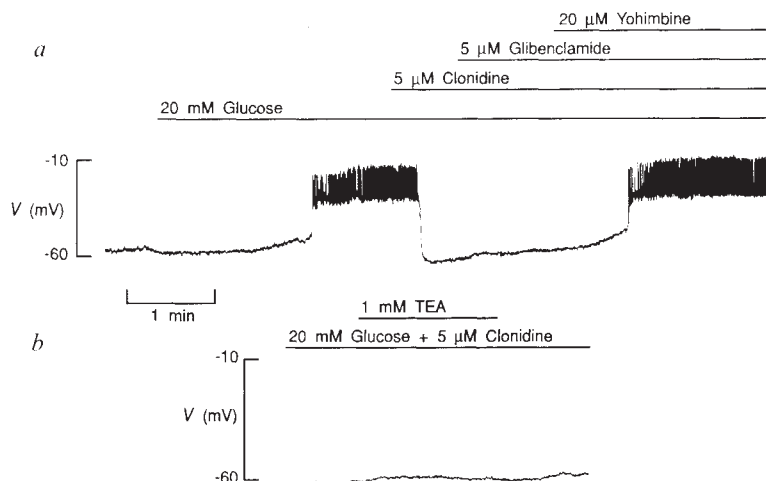


FIG. 3 Repolarizing response to adrenaline does not result from activation of ATP-regulated K^+ channels or Ca^{2+} -activated K^+ channels. Effects of adding glibenclamide (*a*) and TEA (*b*) on clonidine-induced repolarization. *b* is the continuation of *a* after removal of yohimbine.

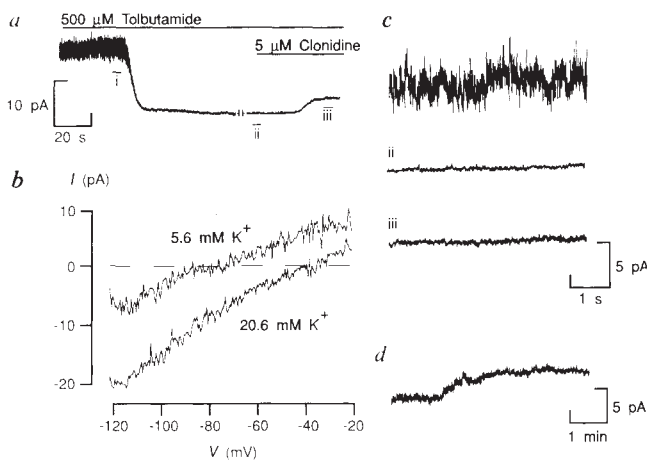


FIG. 4 **a**, Whole-cell voltage-clamp current recorded at a holding potential of -20 mV. **b**, Whole-cell current (I)-voltage (V) relationships for the current activated by clonidine at 5.6 and 20.6 mM $[K^+]_o$. **c**, Current noise observed under basal conditions (trace i), in the presence of tolbutamide (ii) and after the addition of clonidine (iii). Traces were taken from **a** as indicated by the horizontal bars underneath the current trace. **d**, Standard whole-cell recording in the presence of 0.5 mM tolbutamide. The pipette solution contained 0.5 mM GTP- γ -S. Current trace starts 150 s after establishment of the whole-cell configuration.

METHODS. I - V relationships were obtained by subtracting the current response to a voltage ramp between -120 and -20 mV (rate: 62 mV s^{-1}) in the absence of clonidine from that observed after addition of the drug. Curves are mean responses from four different cells. The amplitude of the unitary events was estimated by calculating the variance (σ^2) of the observed current during 30 – 60 s. The value of σ^2 observed in the presence of 500 μ M tolbutamide (panel **d**, trace ii) or 20 mM glucose was taken to reflect the background noise and unrelated to channel activity and the variance in excess of that used to estimate the single-channel current amplitude (i). The value of i was derived using the formula $i = \sigma^2 / (1 - P_{open})$ where i is the whole-cell current and P_{open} is the open probability of the channel; it was ascertained that P_{open} was too low to affect the value of i . The bath contained standard extracellular solution supplemented with 5 mM $CoCl_2$. In **d** the pipette contained (in mM) 125 KCl, 30 KOH, 1.5 MgCl $_2$, 10 EGTA, 5 HEPES-KOH (pH 7.15), 3 Mg-ATP and 0.5 Li $_4$ -GTP- γ -S. The current records were filtered at 100 – 200 Hz.

transmitter release in the endocrine part of the pancreas will consequently impair the secretory activity of the B cell by a mechanism not sensitive to sulphonylureas. This may also account for the clinical observation that some diabetic patients do not respond adequately to sulphonylurea treatment. □

A small GTP-binding protein dissociates from synaptic vesicles during exocytosis

Gabriele Fischer von Mollard*, Thomas C. Südhof† & Reinhard Jahn*‡

* Department of Neurochemistry, Max-Planck Institute for Psychiatry, Am Klopferst 18a, D-8033 Martinsried, Germany

† Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, Texas 75235, USA

Low-molecular-weight GTP-binding proteins are strong candidates for regulators of membrane traffic^{1–3}. In yeast, mutations in the *sec4* or *ypt1* genes encoding small GTP-binding proteins inhibit constitutive membrane flow at the plasma membrane or Golgi complex, respectively^{4–6}. It has been suggested that membrane fusion–fission events are regulated by cycling of small GTP-binding proteins between a membrane-bound and free state⁷, but although most of these small proteins are found in both soluble and tightly membrane-bound forms, there is no direct evidence to support such cycling. In rat brain a small GTP-binding protein, rab3A, is exclusively associated with synaptic vesicles, the secretory organelles of nerve terminals^{8,9}. Here we use isolated nerve terminals to study the fate of rab3A during synaptic vesicle exocytosis. We find that rab3A dissociates quantitatively from the vesicle membrane after Ca^{2+} -dependent exocytosis and that this dissociation is partially reversible during recovery after stimulation. These results are direct evidence for an association–dissociation cycle of a small GTP-binding protein during traffic of its host membrane.

In nerve terminals, neurotransmitters are stored in synaptic vesicles and secreted by fusion of these vesicles with the presynaptic plasma membrane^{10,11}. As this fusion process is highly regulated, synaptic vesicles represent a membrane compartment that is essentially stationary under resting conditions but is rapidly mobilized by Ca^{2+} influx when the nerve terminal is excited. Membrane traffic in the synapse can be tightly controlled *in vitro*, making nerve terminals an ideal system with which to test whether small GTP-binding proteins remain associated with their membranes during fusion–fission events. Here we have used synaptosomes (isolated pinched-off nerve terminals) to investigate the localization of rab3A as a function of synaptic vesicle exocytosis.

Synaptosomes were prepared from rat cerebral cortex by differential and Ficoll-gradient centrifugation. The secretory response of the synaptosomal preparation was monitored by following the release of glutamate using an online enzymatic detection system^{12,13}. As shown in Fig. 1, the synaptosomes responded to K^+ -induced depolarization by glutamate release which was Ca^{2+} -dependent. There was a similar response when the Ca^{2+} -ionophore ionomycin was used for stimulation. Functional viability of the synaptosomes was also confirmed by monitoring the membrane potential with the fluorescent dye 3,3-dipropylthiadicarbocyanine. A high membrane potential was detected in the resting state which was reduced by the addition of K^+ ions (data not shown).

To determine how much rab3A is bound to synaptic vesicles in relation to their excitation state, synaptosomes were stimulated and neurotransmitter release monitored as shown in Fig. 1. Samples were diluted rapidly with ice-cold buffer at different time points in order to block membrane traffic. After hypotonic lysis, a crude synaptic vesicle fraction was isolated by differential centrifugation. The amount of rab3A in the vesicles was determined by immunoblotting and compared with other synaptic vesicle membrane proteins. Potassium-induced stimulation of

Received 23 July; accepted 22 October 1990.

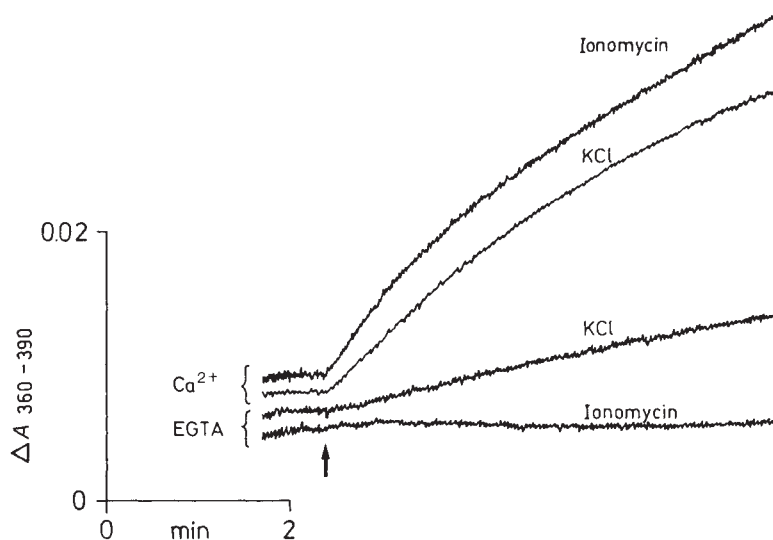
1. Ashcroft, F. M. *A. Rev. Neurosci.* **11**, 97–118 (1988).
2. Henquin, J. C. & Meissner, H. P. *Experientia* **40**, 1043–1052 (1984).
3. Ashcroft, F. M. & Rorsman, P. *Biochem. Soc. Trans.* **18**, 109–111 (1990).
4. Arkhammar, P., Nilsson, T., Rorsman, P. & Berggren, P.-O. *J. Biol. Chem.* **262**, 5448–5454 (1987).
5. Horn, D. R. & Marty, A. *J. gen. Physiol.* **92**, 145–159 (1988).
6. Smith, P. A., Ashcroft, F. M. & Rorsman, P. *FEBS Lett.* **261**, 187–190 (1990).
7. Trube, G., Rorsman, P. & Ohno-Shosaku, T. *Pflügers Arch.* **407**, 493–499 (1986).
8. Schmid-Antomarchi, H., de Weille, J. R., Fosset, M. & Lazdunski, M. *J. Biol. Chem.* **262**, 15840–15844 (1987).
9. Dunne, M. J., Bullett, M.-J. Li, G., Wollheim, C. B. & Petersen, O. H. *EMBO J.* **8**, 413–420 (1989).
10. de Weille, J., Schmid-Antomarchi, H., Fosset, M. & Lazdunski, M. *Proc. natn. Acad. Sci. U.S.A.* **85**, 1312–1316 (1988).
11. Bokvist, K., Rorsman, P. & Smith, P. A. *J. Physiol., Lond.* **423**, 327–342 (1990).
12. Hille, B. *Ionic Channels of Excitable Membrane* 211–216 (Sinauer, Sunderland, 1984).
13. Nilsson, T., Arkhammar, P., Rorsman, P. & Berggren, P.-O. *J. Biol. Chem.* **263**, 1855–1860 (1988).

ACKNOWLEDGEMENTS. This work was supported by the Swedish Medical Research Council, the Swedish Diabetes Association, the Nordisk Insulin Foundation, Swedish Hoechst, the Åke Wibergs Stiftelse, the Magn. Bergvalls Stiftelse, the O.E. och Edla Johanssons Stiftelse, the Lars Hiertas Stiftelse, the Bank of Sweden Tercentenary Foundation and the Svenska Sällskapet för medicinsk forskning.

‡ To whom correspondence should be addressed.

FIG. 1 Ca^{2+} -dependent stimulation of glutamate release by K^{+} -depolarization and ionomycin from isolated synaptosomes. Release was measured spectro-photometrically by following the conversion of glutamate by glutamate dehydrogenase.

METHODS. Synaptosomes were prepared by a modification of the procedure of Nicholls^{19,20}. Cerebral cortices of 2 rats were homogenized in 10 vol ice-cold sucrose (320 mM). The homogenate was cleared of debris by centrifugation at 3,000 *g* for 2 min and a crude synaptosomal fraction was obtained by centrifugation at 14,000 *g* for 12 min. In some experiments this preparation was used without further purification, otherwise the pellet was resuspended in 5 ml sucrose (320 mM) and loaded on top of a discontinuous Ficoll gradient consisting of 13% (4 ml), 9% (1 ml), 5% (4 ml) of Ficoll 400 (all w/v) in 320 mM sucrose, 5 mM HEPES, pH 7.4. The gradient was spun for 35 min at 22,500 r.p.m. in a Beckman SW41 rotor. Synaptosomes enriched in the 9% layer were removed, diluted with 3 vol standard buffer (NaCl, 140 mM; KCl, 5 mM; sodium HEPES, pH 7.4, 20 mM; NaHCO_3 , 5 mM; Na_2HPO_4 , 1.2 mM; MgCl_2 , 1 mM; glucose, 10 mM), fractionated and collected by centrifugation at 14,000 *g* for 12 min. Synaptosomes were stored as pellets on ice before use. For the assay of glutamate release, 1 mg synaptosomal protein was resuspended in 1 ml of standard buffer and incubated under stirring for 15 min at 37 °C. Either CaCl_2 (1.3 mM) or EGTA (0.5 mM) was then added with glutamate dehydrogenase (Sigma type II, 34 U) and 1 mM NADP^{12,13}. After further incubation for ~5 min, additions were made as



indicated (arrow; ionomycin, 10 μM , and KCl, 50 mM, final concentrations) and the incubation extended for 5 min more. Generation of NADPH was monitored by absorbance (*A*) at 360 nm using an Aminco DW 2000 dual wavelength spectrophotometer, using 390 nm as the reference wavelength.

neurotransmitter release led to a large decrease in rab3A in the vesicle fraction (Fig. 2a). The content of both synaptophysin, an integral membrane protein of synaptic vesicles^{14,15}, and of total protein remained constant, demonstrating that the reduction of rab3A is not just due to a nonspecific loss of vesicle membranes (Fig. 2a).

When synaptosomes were stimulated by K^{+} -depolarization in the presence of EGTA to remove extracellular Ca^{2+} , there was no significant dissociation of rab3A (Fig. 2a). This correlates well with the inhibition of glutamate release in the absence of extracellular Ca^{2+} (Fig. 1). Dissociation of rab3A from synaptic vesicles was also observed when neurotransmitter release was stimulated with the Ca^{2+} ionophore ionomycin (Fig. 2b). Again, EGTA inhibited both neurotransmitter release and rab3A dissociation. These data show that the dissociation of rab3A is specifically correlated with Ca^{2+} -dependent exocytosis. Furthermore, Ca^{2+} alone has no effect on the membrane binding of rab3A in an isolated synaptic vesicle fraction (data not shown), indicating that it is not sufficient to dissociate rab3A from the vesicle membrane.

The time course and the extent of rab3A dissociation that occurs during K^{+} -induced neurotransmitter release was then investigated in more detail. As shown in Fig. 3, the dissociation of rab3A is very rapid, with less than 50% remaining bound after two minutes of stimulation, with further reduction during prolonged stimulation (12% after 10 min). These data show that the dissociation is almost quantitative. The 12% remaining is probably due to damaged synaptosomes that did not respond to stimulation.

We also tested whether the dissociation of rab3A from synaptic vesicles after stimulation was reversible. For this purpose, synaptosomes were depolarized for 2 min with 50 mM K^{+} and then allowed to recover in low- K^{+} incubation buffer. As shown in Fig. 4, rab3A continued to dissociate from synaptic vesicles after repolarization of the synaptosomes. But, after prolonged repolarization, rab3A began to reassociate with synaptic vesicles. After 20 min recovery, the rab3A content reached about 50% of the control value, which is fourfold higher than the value at maximal dissociation. The fact that complete reassociation could not be achieved is probably due to partial damage of the synaptosomes (a drop in ATP level for example) caused by the harsh

stimulation conditions and centrifugation and resuspension steps necessary for buffer exchange. This is also reflected by the weak secretory response of these synaptosomes to stimulation (data not shown).

These data indicate that rab3A, a low-molecular-weight GTP-binding protein specific for synaptic vesicles, reversibly dissociates from synaptic vesicles in correlation with exocytosis. This dissociation may occur either during the exocytotic event itself or in a step immediately following exocytosis. After stimulation

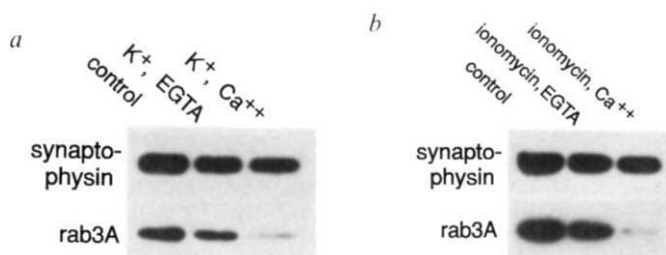


FIG. 2 Rab3A dissociates from a vesicle fraction upon stimulation of synaptosomes in a Ca^{2+} -dependent manner. Transmitter release was stimulated by either K^{+} -depolarization (5 min, 50 mM; a) or ionomycin (5 min, 10 μM ; b). For quantitation, 10 μg protein was analysed by immunoblotting for rab3A and synaptophysin, an integral membrane protein of synaptic vesicles. Note that there is no dissociation in the presence of EGTA.

METHODS. Synaptosomes (2 mg) were resuspended in standard buffer (1.2 ml) and incubated under stirring for 15 min at 37 °C. Addition of CaCl_2 or EGTA and stimulation by K^{+} or ionomycin was as described in the legend to Fig. 1. Control samples were removed and processed immediately before addition of the stimulants. At the end of the incubation, samples were diluted with 6 ml ice-cold standard buffer and centrifuged for 10 min at 12,000 *g*. All subsequent steps were at 4 °C. Pellets resuspended in 0.3 ml standard buffer were lysed by addition of 2.7 ml H_2O , followed by rapid homogenization ($6 \times$ at 2,000 r.p.m. in a glass teflon homogenizer). Electron microscopy after homogenization showed that virtually all synaptosomes were ruptured under these conditions, but without extensive fragmentation of the presynaptic membrane (data not shown). Samples were centrifuged for 10 min at 12,000 *g*. A crude synaptic vesicle fraction was isolated from the supernatant by centrifugation for 20 min at 70,000 r.p.m. in a Beckman T1 100.3 rotor. Pellets were analysed by SDS-PAGE and immunoblotting⁸.

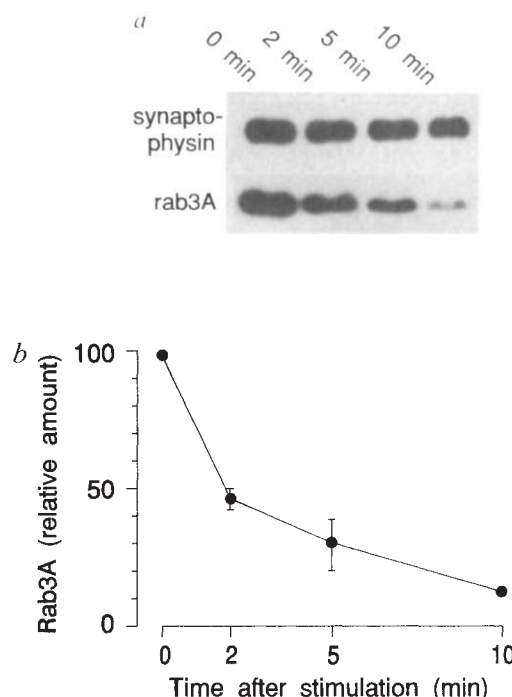


FIG. 3 Time course of rab3A dissociation from crude synaptic vesicles during stimulation of synaptosomes by 50 mM KCl. *a*, Immunoblot analysis of the crude vesicle fraction for synaptophysin and rab3A (times indicated are time intervals after stimulation). *b*, Quantitative determination of rab3A, normalized to the amount of the vesicle marker synaptophysin (0 min represented by 100). With the exception of the 10-min value, all values are means of 2–3 independent experiments (bar indicates range). See legends to Figs 1 and 2 for details. Quantitation was by immunoblotting using ^{125}I -labelled protein A as detection system and γ -radiation counting.

of synaptosomes, a new synaptic vesicle-derived membrane population appears which is devoid of rab3A. We assume that this membrane population represents vesicles that have been retrieved from the plasma membrane following exocytosis. It is believed that the retrieval of functional synaptic vesicles after their fusion with the plasma membrane involves the stepwise formation of coated pits, coated vesicles, an endosome-like compartment, and then budding of new vesicles which are reloaded with neurotransmitter^{10,11,16}. We are not yet able to say at which of these steps rab3A dissociates from synaptic vesicles and when it reassociates. But the dissociation of rab3A is clearly dependent on the exocytotic fusion of the vesicles and not on the rise in intracellular Ca^{2+} alone.

The molecular events underlying the membrane binding and dissociation of rab3A remain to be established. Rab3A is anchored in the membrane by a post-translational covalent modification. As a result, its properties are similar to those of

integral membrane proteins in isolated synaptic vesicles⁸. It can be dissociated from membranes *in vitro*, however, by the GDP-dissociation inhibitor, an enzyme which is specific for rab3 containing bound GDP^{17,18}. Therefore two steps may be necessary for rab3A dissociation: first, a trigger event (GTP hydrolysis) and second, the interaction with the GDP-dissociation inhibitor which catalyses the dissociation of the GDP-bound form of rab3A from the vesicle membrane. Our results indicate that these events are regulated in parallel with exocytosis.

Our results provide direct evidence for the reversible dissociation of a small GTP-binding protein from membranes participating in fusion–fission events. They support the model for the action of small GTP-binding proteins that requires a cycle between membrane-bound and free states for the regulation of membrane traffic. They also show that the synaptosome as a regulated secretory system is ideal for studying this cycle and suggest a specific function for rab3A in directing synaptic vesicles during exocytosis. □

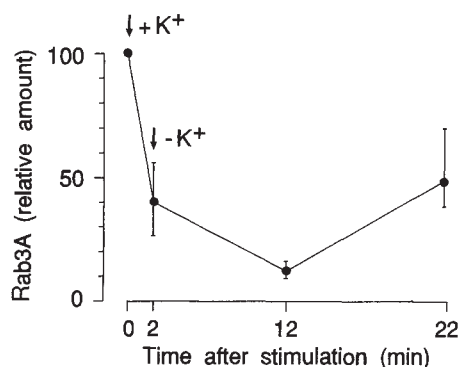


FIG. 4 The dissociation of rab3A is partially reversible. See Fig. 3 for details. Two min after addition of KCl, synaptosomes were spun at room temperature for 90 s at 12,000g in a microfuge and immediately resuspended in 1.2 ml prewarmed standard buffer (low K^+ , no Ca^{2+}). The incubation was continued at 37 °C for the time indicated and processed as described in the legend to Fig. 2. All values are means of 3–4 independent experiments; bar indicates range.

Received 30 August; accepted 22 October 1990.

- Bourne, H. R. *Cell* **53**, 669–671 (1988).
- Balch, W. E. *J. biol. Chem.* **264**, 16965–16968 (1989).
- Hall, A. *Science* **249**, 635–640 (1990).
- Salminen, A. & Novick, P. *J. Cell* **49**, 527–538 (1987).
- Segev, N., Mulholland, J. & Botstein, D. *Cell* **52**, 915–924 (1988).
- Schmitt, H. D., Puzicha, M. & Gallwitz, D. *Cell* **53**, 635–647 (1988).
- Walworth, N. C., Goud, B., Kabacell, A. K. & Novick, P. *EMBO J.* **8**, 1685–1693 (1989).
- Fischer von Mollard, G. *et al. Proc. natn. Acad. Sci. U.S.A.* **87**, 1988–1992 (1990).
- Mizoguchi, A. *et al. J. biol. Chem.* **265**, 11872–11879 (1990).
- Kelly, R. B. *Neuron* **1**, 431–438 (1988).
- De Camilli, P. & Jahn, R. *A. Rev. Physiol.* **52**, 625–645 (1990).
- Nicholls, D. G. & Sihra, T. S. *Nature* **321**, 772–773 (1986).
- Nicholls, D. G., Sihra, T. S. & Sanchez-Prieto, J. *J. Neurochem.* **49**, 50–57 (1987).
- Jahn, R., Schiebler, W., Ouimet, C. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4137–4141 (1985).
- Wiedenmann, B. & Franke, W. W. *Cell* **41**, 1017–1028 (1985).
- Heuser, J. E. & Reese, T. S. *J. Cell Biol.* **57**, 315–344 (1973).
- Sasaki, T. *et al. J. biol. Chem.* **265**, 2333–2337 (1990).
- Araki, S., Kikuchi, A., Hata, Y., Isomura, M. & Takai, Y. *J. biol. Chem.* **265**, 13007–13015 (1990).
- Nicholls, D. G. *Biochem. J.* **170**, 511–522 (1990).
- McMahon, H. T. & Nicholls, D. G. *J. Neurochem.* (in the press).

ACKNOWLEDGEMENTS. We thank D. G. Nicholls (Dundee) for unpublished protocols for the isolation of synaptosomes and for help with the synaptosomal preparation and release assay. This work was supported by a grant from the Deutsche Forschungsgemeinschaft (R.J.) and a NATO travel grant (R.J. and T.C.S.).

Localization of the X inactivation centre on the human X chromosome in Xq13

Carolyn J. Brown*, Ronald G. Lafreniere*, Vicki E. Powers*, Gianfranco Sebastio†, Andrea Ballabio‡, Anjana L. Pettigrew‡, David H. Ledbetter‡, Elaine Levy§, Ian W. Craig§ & Huntington F. Willard*¶

* Department of Genetics, Stanford University School of Medicine, Stanford, California 94305, USA

† Department of Pediatrics, Second Faculty of Medicine, University of Naples, Naples, Italy

‡ Institute for Molecular Genetics, Baylor College of Medicine, Houston, Texas 77030, USA

§ Genetics Laboratory, University of Oxford, Oxford OX3 9DU, UK

¶ To whom correspondence should be addressed

X-CHROMOSOME inactivation results in the strictly *cis*-limited inactivation of many but not all genes on one of the two X chromosomes during early development in somatic cells of mammalian females¹. One feature of virtually all models of X inactivation is the existence of an X-inactivation centre (*XIC*) required in *cis* for inactivation to occur²⁻⁵. This concept predicts that all structurally abnormal X chromosomes capable of being inactivated have in common a defineable region of the X chromosome⁶⁻⁸. Here we report an analysis of several such rearranged human X chromosomes and define a minimal region of overlap. The results are consistent with models invoking a single *XIC* and provide a molecular foothold for cloning and analysing the *XIC* region. One of the markers that defines this region is the *XIST* gene⁹, which is expressed specifically from inactive, but not active, X

chromosomes. The localization of the *XIST* gene to the *XIC* region on the human X chromosome implicates *XIST* in some aspect of X inactivation.

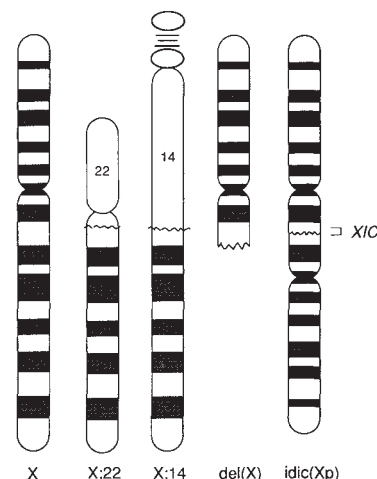
Previous cytogenetic analyses of rearranged X chromosomes that are subject to inactivation have clearly ruled out a requirement for either the short arm or the distal long arm of the X chromosome in X inactivation, thus delimiting the possible location of the *XIC* to a region in the proximal long arm, which is present in all cases of structurally abnormal, inactive X's examined to date⁶⁻⁸. To extend these cytogenetic observations to the molecular level, we have isolated a number of X-chromosome rearrangements with breakpoints in the proximal long arm in rodent-human somatic cell hybrids. The abnormal X chromosomes used in this study, the somatic cell hybrids containing them, and their X-inactivation status¹⁰⁻¹⁵ are shown in Fig. 1.

DNA from hybrid cell lines containing these structurally abnormal X chromosomes was analysed using probes for DNA markers previously known to map to the Xq11-Xq21 region¹⁶. As controls, all probes were hybridized to human, mouse, and hamster genomic DNAs, to DNA from a hybrid containing a normal human X chromosome as its only human component, and to DNA from a hybrid (A63-1A) containing essentially the entire long arm of the X (ref. 17; Fig. 2a). Results of these Southern blot hybridizations are shown in Fig. 2b. The region from the A63 breakpoint in Xq11 to below the 81 deletion breakpoint in Xq13 can be subdivided into five intervals, as shown in Fig. 2a. Considering all of the data, the order of loci in Xq13 is: DXS132-DXS227-*XIST*, DXS128-DXS171, DXS56, *PGK1*-DXS72 (Fig. 2a).

The localization of the *XIST* gene to this region was independently confirmed by *in situ* hybridization. The 8A11 *XIST* cDNA (see accompanying article⁹) was labelled with [³H]dCTP and hybridized to human metaphase chromosomes. In 92 male metaphases examined, the only significant peak of grains observed (14% of the total and 50% of those on the X) was located in band Xq13, at the interface of bands Xq13 and q21.1 (data not

FIG. 1 Structurally abnormal human X chromosomes, each capable of being inactivated. For reference, a normal human X chromosome is also shown (X). For the X;22 (case 68) and X;14 (case 4) translocations involving band Xq13, the autosomal portion of the translocation chromosomes is indicated in white. The derivative X;22 translocation chromosome, isolated in hybrid A68-2A, is from a female with a balanced translocation with the X breakpoint in proximal band Xq13. The subject has a daughter who inherited the der(22) chromosome (containing the X long arm material), but not the der(X) chromosome. In the daughter, the der(22) chromosome was observed to be late-replicating¹⁰, the principal cytogenetic manifestation of X inactivation^{11,12}. The derivative X;14 chromosome was isolated from a male with both reciprocal products of a X;14 translocation, as well as a second copy of the der(14), chromosome that is late-replicating and inactive¹³. The del(X), which is late-replicating, was observed in a female with a deletion of much of the long arm of one X (ref. 14). The idic(Xp) chromosome (Xpter → Xq13::Xq13 → Xpter) was derived from a female patient (A.G.) with symptoms of Turner's syndrome¹⁵. This patient is mosaic for a 45,X cell line and a second cell line with the late-replicating isodicentric X chromosome¹⁵. Because the X inactivation centre must be located on all abnormal X chromosomes that are subject to inactivation, *XIC* maps distal to the breakpoints in the X;22 and X;14 translocations, but proximal to the breakpoints in the del(X) and the idic(Xp), as indicated to the right. Asterisk, *TIMP* and *POLA* are not expressed, see Fig. 3.

CELL LINE	KARYOTYPE	SOMATIC CELL HYBRID	PORTION OF X INACTIVATED	EVIDENCE FOR INACTIVATION	LOCATION OF <i>XIC</i>
68 (GM4628)	46,X,t(X;22)(q13;p11)	A68-2A	q13→qter	Late-replication ¹³	distal to break
4 (GM0074)	47,Y,t(X;14)(q13;q32)+der(14)mat	W4-1A	q13→qter	Late-replication ¹⁶	distal to break
81	46,X,del(X)(pter→q13)	t81-az1b	pter→q13	Late-replication ¹⁷	proximal to break
A.G.	45,X/46,X,idic(Xp)(pter→q13::q13→pter)	tAG-1Baz1b	pter→q13	Late-replication ¹⁸ Gene inactivation*	proximal to break



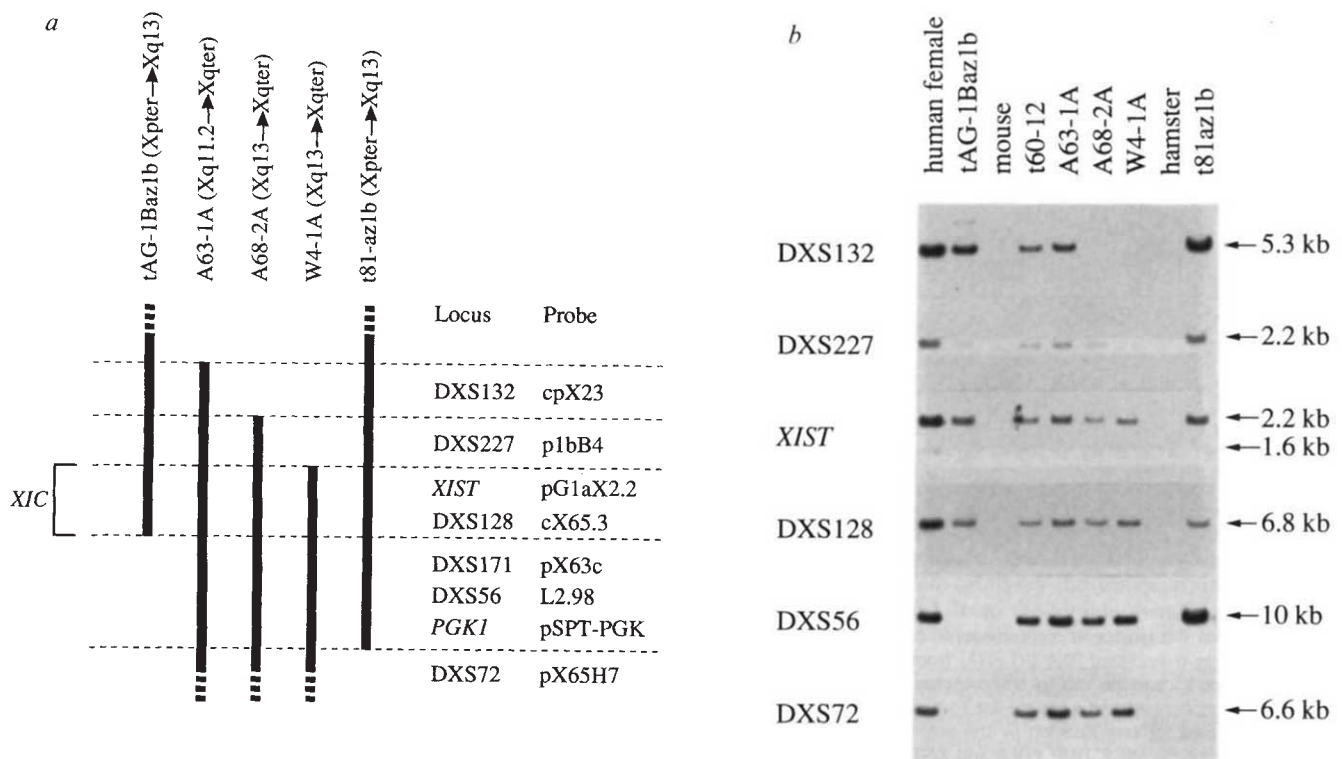


FIG. 2 Regional localization of DNA markers in the proximal long arm of the human X chromosome. **a**, Schematic illustration of the region of overlap in the structurally rearranged X chromosomes contained in hybrids indicated at the top. Locus and probe designations for the DNA markers tested are given at the right. The position of the *XIC* is indicated to the left. **b**, Representative results of Southern hybridization of probes from the Xq13 region to *Hind*III-digested DNA from control human, mouse and hamster lines and from somatic cell hybrids containing different portions of the X chromosome. Sizes of resulting bands are indicated at the right. Figure is a composite from six separate hybridization experiments.

METHODS. Mouse-human hybrids t60-12, A63-1A, A68-2A, and t81az1b and hamster-human hybrid W4-1A have been described^{17-19,27}. Hybrid

tAG-1Baz1b was formed by fusion of lymphoblasts from patient A.G.¹⁵ and mouse cell line tsA1S9, as described^{19,27}. The active X chromosome was removed by back-selection in medium containing 8-azaguanine. Presence of the inactive idic(Xp) and not the normal X was confirmed by extensive karyotyping (by standard methods) and by fluorescence *in situ* hybridization, using a specific alpha satellite DNA probe for the centromere of the X chromosome³². Southern hybridization with ³²P-labelled probes was performed by standard methods^{18,19,22}. Full details on probes can be found in ref. 16. For the *XIST* gene, the pG1aX2.2 probe is a 2.2-kb *Xba*I genomic fragment at the 5' end of the known transcribed sequences and was isolated from one of the genomic phages described in ref. 9.

shown). No significant signal was detected on the Y chromosome or on any of the autosomes, consistent with *XIST* being a single-copy locus.

By combining the map in Fig. 2a with the knowledge of which portion of the X in these cell lines is able to be inactivated, it is possible to identify the region containing the *XIC* (Fig. 1). Of the probes tested, only the *XIST* and DXS128 loci are common to all of the rearranged X chromosomes subject to inactivation. Thus, as predicted by the single *XIC* model of X inactivation, there is a single region within Xq13 which is present on all inactive X chromosomes studied. In conjunction with other mapping data using these hybrids (refs 16, 18, 19, and data not shown), the order of loci around the *XIC* on the human X chromosome is: centromere-*AR-CCG1-PHKA-XIC-PGK1*-telomere. As our data specifically demonstrate that the *XIC* region lies proximal to *PGK1*, a locus previously associated with *XIC* in humans or with the homologous *Xic* (or *Xce*) locus (loci) in mouse^{20,21}, this information should also help guide efforts to localize the murine X-inactivation centre.

The proximal limit for the *XIC* region is currently defined by the breakpoint in a X;14 translocation, in which the der(14) chromosome, containing the Xq13→qter portion of the X chromosome, is inactivated¹³ (Fig. 1). The distal limit for the *XIC* region is defined by the breakpoint in a late-replicating isodicentric X (abbreviated idic(Xp))¹⁵. This idic(Xp) chromosome does not contain most of the region of the long arm present

in the inactivated der(14) chromosome (Fig. 1). Thus, providing that the idic(Xp) chromosome is inactive, *XIC* must lie within the small region of overlap in Xq13 between the der(14) and idic(Xp) chromosomes, a region defined by the loci DXS128 and *XIST* (Fig. 2).

To establish that this idic(Xp) chromosome is subject to inactivation, we isolated this chromosome (without the normal X present) in a mouse-human somatic cell hybrid for analysis of transcription of two X-linked genes known to be subject to inactivation^{22,23}. As shown in Fig. 3, neither the *TIMP* gene nor the *POLA* gene is expressed in the idic(Xp) hybrid. These data demonstrate that the idic(Xp) chromosome is inactive and so contains the *XIC*. As shown in Fig. 2b, the *XIST* gene is present on this chromosome and, as expected for an inactive X chromosome, *XIST* is expressed in idic(Xp) hybrid cells (Fig. 3).

The concept of a single X-inactivation centre has long been a part of most X-inactivation models^{2,3,24}. At least two key events in X inactivation have been proposed to involve the *XIC*. First, the *XIC* must be marked before inactivation as a means of distinguishing between which X remains active and which X becomes inactive^{24,25}. Second, an inactivation signal which leads to the *cis*-limited transcriptional inactivation of genes located on either side of *XIC*, but which must be capable of 'skipping' over portions of the chromosome containing genes that escape inactivation²⁶⁻²⁸, has been suggested to spread from this site²⁸⁻³¹. The data reported here provide support for the existence of a

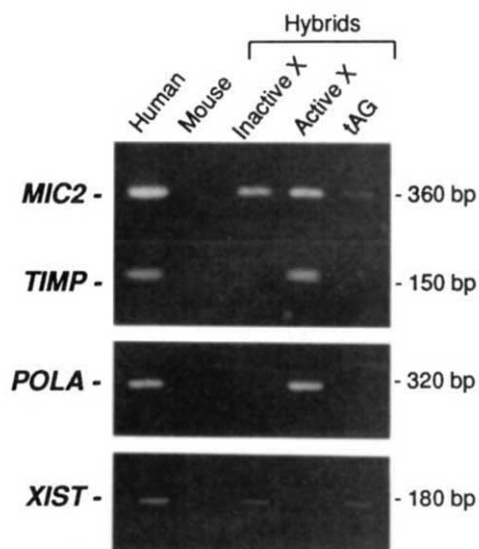


FIG. 3 Evidence that the isodicentric chromosome of A.G. is subject to X inactivation. Reverse-transcribed RNA (RT-RNA) from somatic cell hybrids containing an X_i and X_a , and the idic(Xp) chromosome (TAG), as well as from a human female lymphoblast and a mouse line (the parental mouse line for the hybrids) was used for amplification by the polymerase chain reaction (PCR) with primers for the *MIC2*, *TIMP*, *POLA*, and *XIST* genes. The *TIMP* and *POLA* genes have previously been demonstrated to be subject to X inactivation, while the *MIC2* gene is known to escape X inactivation^{22,23}. The *MIC2* primers amplified a 360-bp product for all RT-RNAs except for the mouse. The *TIMP* primers were used in a duplex reaction with the *MIC2* primers. They amplified a 150-bp product in the human RT-RNA and the X_a hybrid RT-RNA. The *POLA* primers also only amplified their 320-bp product in the female and X_a hybrid RT-RNA. The 180-bp *XIST* product was amplified from the female, the X_i hybrid, and the TAG hybrid. Therefore, the TAG hybrid behaved identically to the X_i hybrid.

METHODS. RT-PCR was performed as described^{9,22}, except that 10 times the input RT-RNA was used for the *POLA* primers. The *MIC2* and *TIMP* primers have been described²². The *POLA* primers are 5'-TGGCCATTTCAT-CACCCAGT-3' and 5'-ACTGCCATACTGAAATACAT-3' which amplify a predicted 320-bp product³³. The *XIST* primers were 1→ and 2← as described in the accompanying article⁹. The X_a and X_i hybrids used were as described^{9,22}.

single *XIC* on the human X chromosome and, accordingly, should significantly refine efforts to clone and analyse this locus. That the *XIC* region is coincident (at the current level of mapping resolution) with the location of the *XIST* gene⁹, whose expression pattern is specifically and uniquely affected by the inactivation status of the X chromosome on which it lies, strongly implicates *XIST* in some aspect of the X inactivation process. □

Received 16 October; accepted 13 November 1990.

1. Lyon, M. F. *Nature* **190**, 372-373 (1961).
2. Lyon, M. F. *Nature new Biol.* **232**, 229-232 (1971).
3. Russell, L. B. *Science* **140**, 976-978 (1963).
4. Gartler, S. M. & Riggs, A. D. *Rev. Genet.* **17**, 155-190 (1983).
5. Cattanach, B. M. *A. Rev. Genet.* **9**, 1-18 (1975).
6. Mattei, M. G., Mattei, J. F., Vidal, I. & Giraud, F. *Hum. Genet.* **56**, 401-408 (1981).
7. Summitt, R. L., Tipton, R. E., Wilroy, R. S. Jr, Martens, P. & Phelan, J. P. *Birth Defects: Orig. Art. Ser.* **14**, 219-247 (1978).
8. Therman, E., Sarto, G. E. & Patau, K. *Chromosoma* **44**, 361-366 (1974).
9. Brown, C. J. *et al. Nature* **349**, 38-44 (1991).
10. Buhler, E. M., Jurik, L. P., Voyame, M. & Buhler, U. K. *Nature* **265**, 142-144 (1977).
11. Morishima, A., Grumbach, M. M. & Taylor, J. H. *Proc. natn. Acad. Sci. U.S.A.* **48**, 756-763 (1962).
12. Latt, S. A. *Expt Cell Res.* **86**, 412-415 (1974).
13. Allderdice, P. W., Miller, O. J., Miller, D. W. & Klinger, H. P. *Am. J. med. Genet.* **2**, 233-240 (1978).
14. Tantravahi, U. *et al. Hum. Genet.* **64**, 33-38 (1983).
15. Pettigrew, A. L., McCabe, E. R. B., Elder, F. F. B. & Ledbetter, D. H. *Am. J. hum. Genet.* **43**, A118 (1988).
16. Mandel, J. L. *et al. Cytogenet. Cell Genet.* **51**, 384-437 (1989).

17. Mahtani, M. M. & Willard, H. F. *Genomics* **2**, 294-301 (1988).
18. Brown, C. J., Powers, V. E., Monroe, D. M., Sheinin, R. & Willard, H. F. *Somat. Cell molec. Genet.* **15**, 93-96 (1989).
19. Brown, C. J. *et al. Am. J. hum. Genet.* **44**, 264-269 (1989).
20. Keer, J. T. *et al. Genomics* **7**, 566-572 (1990).
21. Johnston, P. G. & Cattanach, B. M. *Genet. Res.* **37**, 151-160 (1981).
22. Brown, C. J., Fienniken, A., Williams, B. & Willard, H. F. *Nucleic Acids Res.* **18**, 4191-4195 (1990).
23. Wang, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 5270-5274 (1985).
24. Brown, S. W. & Chandra, H. S. *Proc. natn. Acad. Sci. U.S.A.* **70**, 195-199 (1973).
25. Rastan, S. J. *Embryol. exp. Morph.* **78**, 1-22 (1983).
26. Schneider-Gadicke, A., Beer-Romero, P., Brown, L. G., Nussbaum, R. & Page, D. C. *Cell* **57**, 1247-1258 (1989).
27. Brown, C. J. & Willard, H. F. *Am. J. hum. Genet.* **45**, 273-279 (1990).
28. Mohandas, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 4954-4958 (1987).
29. Rastan, S. & Robertson, E. J. *J. Embryol. exp. Morph.* **90**, 379-388 (1985).
30. Lyon, M. F. *et al. J. Embryol. exp. Morph.* **97**, 75-85 (1986).
31. Riggs, A. D. *Aust. J. Zool.* **37**, 419-441 (1990).
32. Sharp, C. B., Bedford, H. M. & Willard, H. F. *Hum. Genet.* **85**, 330-336 (1990).
33. Wong, S. W. *et al. EMBO J.* **7**, 37-47 (1988).

ACKNOWLEDGEMENTS. We thank P. L. Pearson, K. E. Davies and B. White for DNA probes, and K. E. Davies, J. Rupert and L. Carrell for discussions. This work was supported by the NIH.

The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the *Tme* locus

D. P. Barlow*, R. Stöger*, B. G. Herrmann†, K. Saito‡ & N. Schweifer*

* Research Institute for Molecular Pathology, Dr Bohr-Gasse 7, A-1030 Vienna, Austria
† MPI für Entwicklungsbiologie, Spemannstrasse 35, D-7400 Tübingen, Germany
‡ Vanderbilt University, Nashville, Tennessee 37232-2172, USA

T-ASSOCIATED maternal effect (*Tme*) is the only known maternal-effect mutation in the mouse^{1,2}. The defect is nuclear-encoded³ and embryos that inherit a deletion of the *Tme* locus from their mother die at day 15 of gestation⁴. There are many genomically imprinted regions known in the mouse genome^{5,6}, but so far no imprinted genes have been cloned. The *Tme* locus is absent in two chromosome-17 deletion mutants, *T^{hp}* and the *t^{Lub2}*, and its position has been localized using these deletions to a 1-cM region⁷⁻¹⁰. We report here that the genes for insulin-like growth factor type-2 receptor (*Igf2r*) and mitochondrial superoxide dismutase-2 (*Sod-2*) are absent from both deletions. Probes for these genes and for plasminogen (*Plg*) and T-complex peptide 1 (*Tcp-1*) were used in pulsed-field gel mapping to show that *Tme* must lie within a region of 800-1,100 kb. We also demonstrate that embryos express *Igf2r* only from the maternal chromosome, and that *Tcp-1*, *Plg* and *Sod-2* are expressed from both chromosomes. Therefore *Igf2r* is imprinted and closely linked or identical to *Tme*.

The position of the *T^{hp}* and *t^{Lub2}* deletions relative to cloned DNA markers^{1,8-10} is shown in Fig. 1. Of these marker loci, only *Tcp-1* is deleted in the *t^{Lub2}* chromosome, so the distance between the flanking marker loci *D17Rp17* and *D17Leh66D* gives the closest approximation of the limits of the region containing the *Tme* gene. Cumulative mapping data indicate that these markers are separated by at least one centimorgan⁷.

Tcp-1 (ref. 11) and a gene from the *D17Leh66D* locus, *Tcp-10* (ref. 12), map in the human to chromosome 6q21-27 in close linkage to the plasminogen, insulin-like growth factor type-2 receptor and superoxide dismutase-2 loci¹³. These last three genes have been assigned to mouse chromosome 17 (refs 14-16) but not mapped with respect to the *Tme* locus. Figure 2 shows the mapping of these three genes in the *t^{Lub2}* deletion chromosome and the *Ti^{Orl}* chromosome that is presumed to contain a duplication of the *Tme* locus (see refs 10 and 17 for a description

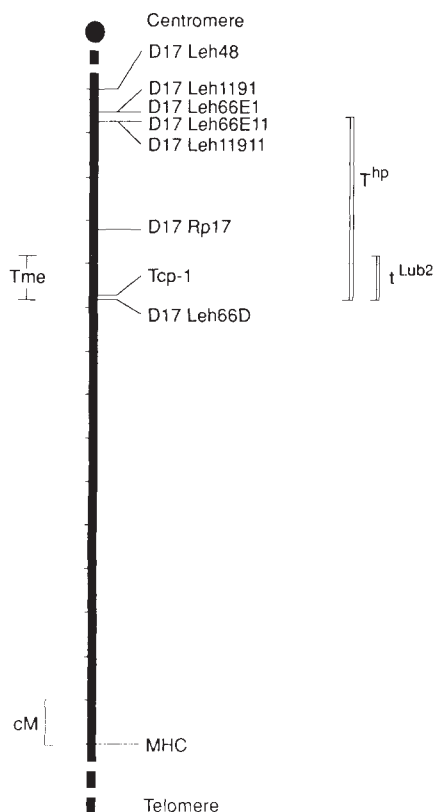
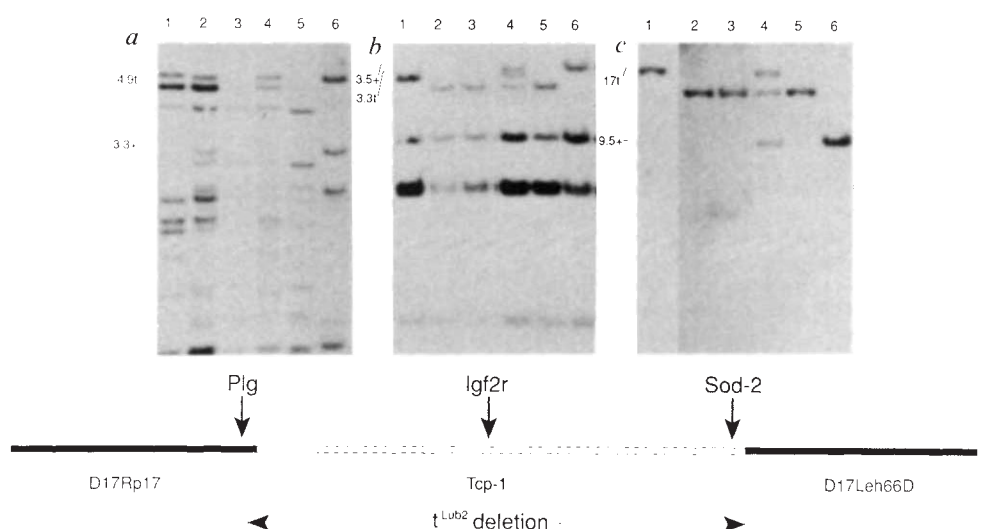


FIG. 1 Genetic map⁷ of the proximal part of mouse chromosome 17 showing the position of the T^{hp} and the t^{Lub2} deletions used to define the limits of the region containing the Tme locus^{9,10}.

of the structure and origin of the t haplotype and partial t haplotype chromosomes). Our results show that the *Igf2r* and *Sod-2* loci are deleted in t^{Lub2} and duplicated in Tt^{Ori} . The locus for *Plg* is present in two copies in t^{Lub2} , one of which is partially deleted and present once in Tt^{Ori} . Subsequent work has identified the *Sod-2* and *D17Leh66D* loci as residing on the same cosmid clone and thus the distal end of the t^{Lub2} deletion can be localized to within 45 kilobases (kb) of the *Sod-2* locus (K.S. and D.P.B., data not shown). In addition, further study of the *Plg* locus in the t^{Lub2} and Tt^{Ori} chromosomes⁹ has shown that this locus forms the proximal breakpoint of the t^{Lub2} deletion (N.S. and D.P.B., data not shown). Thus the region that contains the Tme locus can be exactly defined as lying between the *Plg* and *Sod-2* loci. Genetic mapping has separated these two loci¹³, but the estimated genetic distance of ~1 cM does not differ from that estimated for *D17Rp17* and *D17Leh66D*. Recombination frequencies are not an accurate measure of physical distance and we therefore constructed a long-range restriction map using DNA separated by pulsed-field gel electrophoresis¹⁸ in order to size the region containing the Tme locus. Figure 3 shows a long-range restriction map including the *D17Rp17* and the *D17Leh66D* loci and we have used this to estimate the separation of the *Plg* and *Sod-2* loci as 0.8–1.1 megabases (Mb).

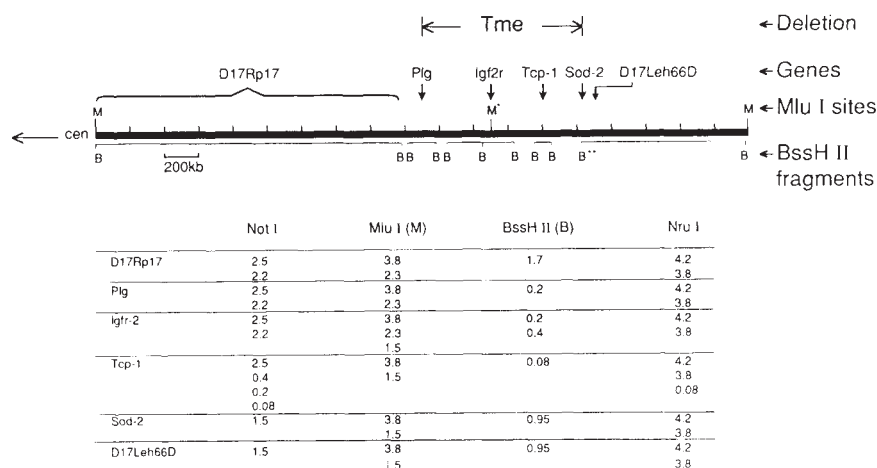
These results and the previous mapping of the *Tcp-1* gene⁹ thus assigns four genes to the region known to contain the Tme locus. Next we decided to test these four genes for evidence of imprinting. We reasoned that the Tme gene would only be functionally expressed from the maternal chromosome. Therefore the Tme embryonic-lethal phenotype could be explained by the absence or alteration of the Tme transcript when the embryo inherits a deletion of the maternal locus and contains only the paternal Tme locus. The normal transcript would be present in the converse situation, when the embryo inherits a deletion of the paternal locus and contains only the maternal locus. Non-imprinted genes would be expressed from both

FIG. 2 Mapping using deletion chromosomes. *a*, *TaqI*-digested mouse genomic DNAs hybridized with a mouse plasminogen cDNA clone (MP33B; provided by S. J. F. Degan¹⁴); *b*, *PvuII*-digested DNAs hybridized with an insulin-like growth factor type-2 receptor cDNA clone (mouse partial cDNA; provided by A. Ullrich); *c*, *BamHI*-digested DNAs hybridized with a mitochondrial superoxide dismutase-2 mouse cDNA clone (provided by G. I. Bell¹⁵). The lane order is the same for all three panels: lane 1, t^{w12}/t^{w12} ; lane 2, $t^{Lub2}/M.spretus$; lane 3, $T^{hp}/M.spretus$; lane 4, $Tt^{Ori}/M.spretus$; lane 5, *M.spretus*; lane 6, *C57Bl/6*. The three deletion chromosomes (T^{hp} , t^{Lub2} and Tt^{Ori})^{9,10} were bred over the *spretus* chromosome because of the ease with which polymorphisms can be found between *Mus spretus* and *Mus musculus*; all other chromosomes in the panels have a *Mus musculus* origin, referred to as + or wild-type. Lanes 1, 5 and 6 are control lanes that identify the specific restriction fragments for these three DNA types. The *Plg* panel shows that a 4.9-kb t -specific fragment is present in t^{Lub2} and Tt^{Ori} , and a 3.3-kb + specific fragment is absent in T^{hp} and Tt^{Ori} but present in t^{Lub2} . These and other data (not shown here) place *Plg* at the proximal border of the t^{Lub2} deletion. The *Igf2r* panel shows that a 3.3-kb t -specific fragment, and a 3.5-kb + specific fragment are absent in t^{Lub2} and T^{hp} respectively, but are both present in Tt^{Ori} . The Tt^{Ori} chromosome contains a duplication of the region that is deleted in the t^{Lub2} chromosome, including the Tme locus⁹. The data here place the *Igf2r*



locus within the t^{Lub2} deletion and within the duplicated region in Tt^{Ori} . The *Sod-2* panels shows that a 17.0-kb t -specific fragment and a 9.5-kb + specific fragment, are absent in T^{hp} and t^{Lub2} but are both present in Tt^{Ori} . These data place the *Sod-2* locus within the t^{Lub2} deletion and within the region duplicated in Tt^{Ori} , further analysis of cosmid clones (data not shown) places the *Sod-2* locus at the distal border of the t^{Lub2} deletion. The positions of the *Plg*, *Igf2r* and *Sod-2* loci relative to the t^{Lub2} deletion and existing markers for this region^{9,10} are shown underneath. Preparation of genomic DNA, enzyme digestion, DNA blots and hybridizations were all achieved using standard techniques¹⁰.

FIG. 3 Physical mapping using pulsed-field gels. Genomic DNA from the BALB/c inbred strain was prepared in agarose blocks, digested with rare cutter enzymes and DNA fragments in the 1–5 Mb range were separated by pulsed-field gel electrophoresis. Fragment sizes obtained with four enzymes are given. The map was constructed by sequential hybridization of one filter with the probes listed, and is centred on two neighbouring *Mlu*I fragments. The *Igf2r* gene spans the central *Mlu*I (M^*) site, which was partially cut in genomic DNA. *D17Rp17* is located on a 1.7-Mb *Bss*HI fragment contained within the 2.3-Mb *Mlu*I fragment. *Sod-2* spans a *Bss*HI site (B^{**}) and is located on a 0.95-Mb fragment which is contained within the 1.5-Mb *Mlu*I fragment. If these two *Bss*HI fragments are positioned at the ends of the two *Mlu*I fragments, then the *Plg* and *Sod-2* loci are separated at most by 1.1 Mb, and this is the maximum size of the region which contains *Tme*. The minimum distance separating these two markers (~ 0.8 Mb) is estimated by summation of the *Bss*HI and *Not*I fragments recognized by *Igf2r* and *Tcp-1* but not by the flanking markers. METHODS. Pulsed-field gel DNA preparation using the spleen from BALB/c female mice, enzyme digestion, and electrophoresis were as described²⁴ (detailed protocols available on request). DNA fragment sizes estimated after pulsed-field gel separation are given in megabases and listing of more than one fragment size indicates partial digestion. The *Mlu*I, *Not*I and *Nru*I fragments were sized on pulsed-field gels separating 1–5 Mb over 5 cm



using *S. pombe* chromosomes as size markers²⁵. The *Bss*HI fragments were sized on gels separating 0.1–1.5 Mb over 12 cm, using *S. cerevisiae* YP148 chromosomes as size markers (P. Heiter, personal communication). Probes: as described in Fig. 2, plus *D17Rp17* (DNA segment Roswell Park 17) provided by R. Elliott²⁶, *Tcp-1* (t-complex protein-1) provided by K. Willison¹¹ and *D17Leh66D* (DNA segment Lehrach 66D) provided by H. Lehrach¹⁰.

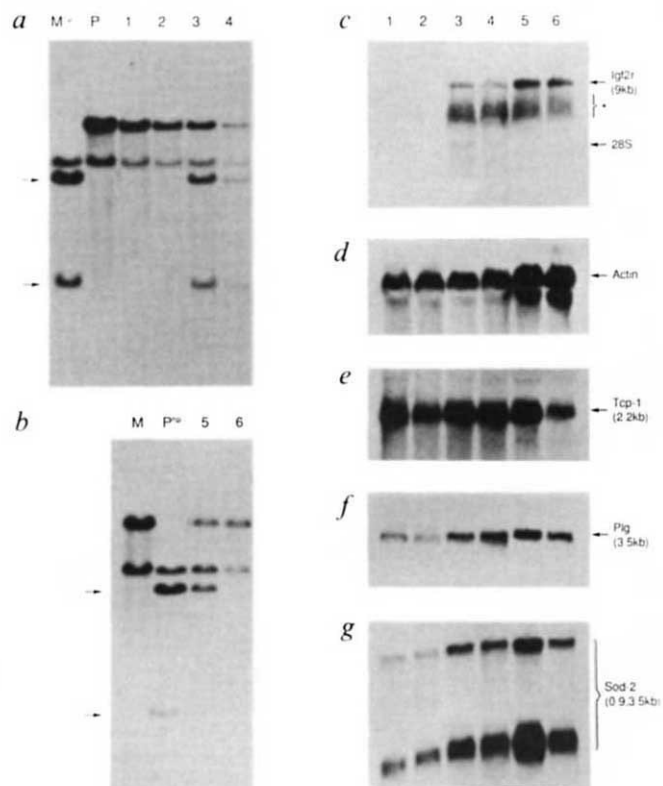
chromosomes in proportion to the number of chromosomes present.

Figure 4 shows the results of an analysis of steady-state messenger RNA levels of the *Igf2r*, *Tcp-1*, *Plg* and *Sod-2* genes in 15-day-old embryos (E15) that inherited a deletion of *Tme*

from their mother (embryos 1 and 2) or their father (embryo 6). The parental and embryo genotypes were identified using a marker from the *D17Rp17* locus (Fig. 4a and b). From the argument presented above, an imprinted gene would show no expression in embryos 1 and 2, but would show expression in

FIG. 4 Genotype (a and b) and mRNA expression analysis (c–g) of E15 embryos. a, Maternal inheritance of the *Tme* deletion. M^{hp} : mother, *Tme*^{hp}/C3H.P; father, BALBc/BALBc. Lanes 1 and 2: embryos of genotype *Tme*^{hp}/BALBc. Lanes 3 and 4: embryos of genotype C3H/BALBc. b, Paternal inheritance of the *Tme* deletion. M: mother, BALBc/BALBc; P^{hp} : father, *Tme*^{hp}/C3H. Lane 5: embryo of genotype C3H/BALBc. Lane 6: embryo of genotype *Tme*^{hp}/BALBc. The absence, in embryos 1, 2 and 6, of the two *Dra*I C3H-specific restriction fragments (indicated by the arrows) shows that they have inherited the *Tme* deletion chromosome. c–g, Messenger RNA expression in embryos 1–6. c, Hybridization of RNA blots using *Igf2r*, and d, hybridization of the filter shown in c with an actin probe. e, f and g. Three separate RNA blots prepared from the same samples as in c, hybridized with *Tcp-1*, *Plg* and *Sod-2* cDNAs respectively. This mRNA analysis shows that expression of *Igf2r* mRNA is absent in embryos that inherit a deletion of the *Tme* locus from their mother (embryos 1 and 2) and compared with their wild-type litter mates (embryos 3 and 4), when equivalent amounts of RNA can be detected using the actin probe (d). Embryos that inherit a deletion of the *Tme* locus from their father (embryo 5) express *Igf2r* at levels similar to wild-type control litter mates (embryo 6). These results have been confirmed in a second cross analysing three maternally derived deletions and two paternally derived deletions (data not shown) and identify *Igf2r* as being expressed from only the maternal chromosome. The expression of *Plg*, *Tcp-1* and *Sod-2* mRNAs are independent of the parental origin of the chromosome, embryos 1 and 2 show hemizygous levels of expression for these three genes compared with wild-type embryos 3 and 4. The same hemizygous levels of expression are seen for these genes when embryo 6 is compared with its littermate control, embryo 5.

METHODS. Embryos were killed at day 15 and separated from extra-embryonic tissue. Embryonic tissue was minced and half was used to prepare DNA¹⁰ and half to prepare RNA²⁷. DNA was digested with *Dra*I to prepare blots which were hybridized with probe Rpb2, a *Bam*HI fragment from a cosmid clone isolated using *D17Rp17* marker (D.P.B., unpublished data). The *D17Rp17* locus is deleted from the *Tme*^{hp} chromosome²⁶. *Tme*^{hp}/C3H male and female mice were provided by E. Eicher. RNA was analysed on formaldehyde gels²⁷ and hybridized as shown. The migration position of 28S ribosomal RNA is indicated (c). *Tme* mutant embryos die at E15 and normally show a generalized body oedema before death, but no other abnormality is evident^{1,2,4}. To preserve the integrity of the tissues, embryos were killed before the onset of visible oedema. However in c, *Igf2r* mRNA is slightly



degraded, producing a band of ~ 9 kb (indicated by arrow) and a smear (*) that extends down to the 28S ribosomal band. This degradation was a feature of many RNA preparations and may be a result of the large size of the mRNA. Hybridization of this filter with actin cDNA (d) shows roughly equal RNA loading for embryos 1, 2, 3 and 4, but embryos 5 and 6 contain more RNA.

embryo 6, at levels comparable to wild-type embryos. Non-imprinted genes would show expression in embryos 1, 2 and 6 that would be equal to half that in wild-type controls. From the RNA blots in Fig. 4, it can be seen that only *Igf2r* shows the predicted expression pattern of an imprinted gene. The *Plg*, *Tcp-1* and *Sod-2* genes all have the predicted expression pattern of non-imprinted genes.

Our results show that the *Igf2r* gene maps to the *Tme* locus and is maternally imprinted. The expression of three other genes that also map to the same region is not apparently influenced by parental origin. Although confirmation of *Igf2r* as the *Tme* gene awaits rescue or inactivation experiments, immediate use can be made of the results described here to examine the role of the *Igf2* receptor in development and the molecular basis of imprinting. The *Tme* mutation is lethal at day 15 of embryogenesis, and embryos are oedematous but do not show any growth defects⁴. If the *Igf2* receptor is the *Tme* gene, this phenotype contrasts with that generated by inactivation of the *Igf2* ligand gene¹⁹ which resulted in small but viable mice. Because the *Igf2* receptor lacks signal-transducing properties and is identical to the mannose-6-phosphate receptor²⁰, its role in mediating the growth-regulating properties of the *Igf2* ligand is not clear but may be better understood by studies with the *Tme* mutant. The molecular basis of genomic imprinting is unknown but we are now in a position to use the *Igf2r* gene to identify the mechanisms through which parental origin can modify gene expression in development. Our immediate goals are to examine the regulatory regions of the *Igf2r* genomic locus for evidence of parental-specific epigenetic modification. The only known type of epigenetic modification in mammalian DNA is cytosine methylation in the CpG dinucleotide²¹. Evidence from several sources indicates that methylation can interfere with gene expression and the availability of a genomic clone will allow us to test this and other possible mechanisms of imprinting. Whether genomic imprinting exists as a specific mechanism to regulate gene expression in development or whether it is a nonspecific mechanism, the consequence of which would prevent asexual reproduction, is still not clear^{22,23}. The identification here of the *Igf2* receptor as an imprinted gene, and the recent report that the *Igf2* ligand also shows parental-specific differences in gene expression¹⁹ mean that molecular tools are now available to answer questions about genomic imprinting. □

Received 3 September; accepted 22 October 1990.

1. Johnson, D. R. *Genetics* **76**, 795-805 (1974).
2. Johnson, D. R. *Genet. Res. Camb.* **24**, 207-213 (1975).
3. McGrath, J. & Solter, D. *Nature* **308**, 550-551 (1984).
4. Alton, A. K. thesis, Cornell Univ. Med. Coll. (1982).
5. Beechey, C. V., Cattanaach, B. M. & Searle, A. G. *Mouse Genome* **87**, 64-65 (1990).
6. Surani, M. A., Reich, W. & Allen, N. D. *Trends Genet.* **4**, 59-62 (1988).
7. Committee for the Mouse Chromosome 17 *Map of Mouse Chromosome 17: First Report, Mammalian Genome* (in the press).
8. Winking, H. & Silver, L. M. *Genetics* **108**, 1013-1020 (1984).
9. Sarvetnick, N. *et al. Genetics* **113**, 723-734 (1986).
10. Herrmann, B. G. *et al. Cell* **44**, 469-476 (1986).
11. Willison, K. *et al. EMBO J.* **6**, 1967-1974 (1987).
12. Bibbins, K. B. *et al. Genomics* **5**, 139-143 (1989).
13. Spence, M. A., Spurr, N. K. & Field, L. L. *Cytogenet. Cell Genet.* **51**, 149-165 (1989).
14. Degan, S. J. F., Bell, S. M., Schaefer, L. A. & Elliott, R. W. *Genomics* (in the press).
15. Laureys, G., Barton, D. E., Ullrich, A. & Francke, U. *Genomics* **3**, 224-229 (1988).
16. Figueroa, F., Vincek, V., Kasahara, M., Bell, G. I. & Klein, J. *Immunogenetics* **28**, 260-264 (1988).
17. Silver, L. M. *A. Rev. Genet.* **19**, 179-208 (1985).
18. Schwartz, D. C. & Cantor, C. R. *Cell* **37**, 67-75 (1984).
19. DeChiara, T., Efstratiadis, A. & Robertson, E. *Nature* **345**, 78-80 (1990).
20. Roth, R. A. *Science* **239**, 1269-1271 (1988).
21. Holiday, R., Monk, M. & Pugh, J. E. *Phil. Trans. R. Soc. B* **236**, 177-338 (1990).
22. Solter, D. *A. Rev. Genet.* **22**, 127-146 (1988).
23. Surani, M. A. *et al. Phil. Trans. R. Soc. B* **326**, 179-187 (1990).
24. Herrmann, B. G., Barlow, D. P. & Lehrach, H. *Cell* **48**, 813-825 (1987).
25. Fan, J.-B. *et al. Nucleic Acids Res.* **17**, 2801-2818 (1988).
26. Mann, E. A., Silver, L. M. & Elliott, R. W. *Genetics* **114**, 993-1006 (1986).
27. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor, 1989).

ACKNOWLEDGEMENTS. We are very grateful to G. I. Bell, S. J. F. Degan and A. Ullrich for the *Sod-2*, *Plg* and *Igf2r* complementary DNA clones. We thank E. Eichler, H. Winking and J.-L. Guenet for *T^{hpo}*, *t^{Lub2}* and *Tt^{Or1}* mice; A. Ullrich for discussions about the *Igf2* receptor; H. Lehrach for advice and encouragement in the early stages of this project; and E. F. Wagner and A. P. Bird for comments on the manuscript.

nature is available in microform.



University Microfilms

International reproduces this publication in microform: microfiche and 16mm or 35mm film. For information about this publication or any of the more than 13,000 titles we offer, complete and mail the coupon to: University Microfilms International, 300 N. Zeeb Road, Ann Arbor, MI 48106. Call us toll-free for an immediate response: 800-521-3044. Or call collect in Michigan, Alaska and Hawaii: 313-761-4700.

☐ Please send information about these titles:

Name _____

Company/Institution _____

Address _____

City _____

State _____ Zip _____

Phone (____) _____

**University
Microfilms
International**

Out with the old, in with the new

Rare-cutting restriction enzymes, a novel method for cloning and packaging DNA fragments and sequence processor software for the Macintosh — new products for the new year.

THE Clinipak **bench-top sharps container** for the disposal of needles, syringes and scalpels reduces the risk of laboratory personnel being exposed to infection or contamination, says Radleys (*Reader Service No. 100*). The compact Clinipak container, which has a five-litre capacity,

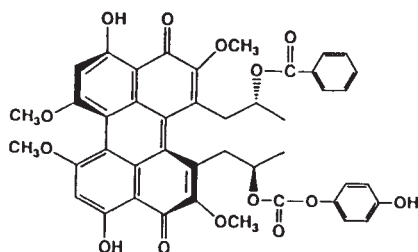


Clinipak sharps disposal unit.

has a device built into the lid that allows the safe removal of a needle from the body of a syringe, without the user having to touch the needle itself. Objects of larger dimensions can be deposited in the container via an elliptical opening, which can be snapped shut after use, ready for disposal and incineration. The container is made of a durable 1-mm thick plastic and is sold in cases of 40 containers.

Assorted reagents

Calphostin C, a newly isolated compound from *Cladosporium cladosporoides*, is a potent and specific **inhibitor of protein kinase C** with an IC_{50} value of 0.05 μ M, says Kamiya Biomedical (*Reader Service No. 101*). The new product is designed to complement Staurosporine, which the company claims is the most specific protein kinase C inhibitor available. Protein



Calphostin C

Calphostin C: protein kinase C inhibitor.

kinase C plays an important role in the regulation of various cellular functions — signal transduction, cellular proliferation and differentiation. Calphostin C is said to inhibit the regulatory domain of protein kinase C, rather than the more tightly conserved catalytic domain. It does not compete with Ca^{2+} or phospholipid binding, but does inhibit 3H PDDBu binding to protein kinase C, says Kamiya.

Transfectam, a new **transfection aid** from IBF biotechnics, is designed for the simple and rapid transfection of cell lines (*Reader Service No. 102*). The reagent is a synthetic cationic lipid that coats plasmidic DNA, which, in turn, promotes binding to cell membranes and incites endocytosis. The transfection process is highly specific and stable, taking an average of 30 minutes, says IBF. Transfectam has been used successfully to transfect HepG2, HeLa, CHO, S49, 3T3, MRC 5 and COS cell lines; the efficiency of the transfection process with CHO and 3T3 has been shown to be up to 1,000 times better than calcium phosphate or liposome techniques, says the company. IBF supplies Transfectam as a sterile dry powder, which, after solubilization, is stable for up to four months at 4 °C.

Molecular market place

Janssen Biochimica offers a whole host of **rare-cutting restriction enzymes** supplied in low concentrations with optimized buffer (*Reader Service No. 103*). The line up of nearly 100 restriction enzymes includes *Aat*II (GACGT/C), *Cl*aI (AT/CGAT), *Not*I (GC/CGCCGC), *Pvu*I (CGAT/CG), *Sac*II (CCGC/GG), *Sal*I (G/TCGAC), *Sma*I (CCC/GGG) and *Xho*I (C/TCGAG). Each vial is supplied with a protocol, a description of the storage and assay buffers, and a surplus quantity of optimized assay buffer (10 × concentration). The restriction enzymes remain stable for 12 months when stored at -20 °C, says Janssen, and are available in high concentrations and bulk quantities on request.

According to Ambion, its new RNase Cocktail, a **mixture of highly purified ribonucleases A and T1**, offers advantages in terms of effectiveness, convenience and purity over the use of RNase A alone (*Reader Service No. 104*). The RNase T1 is purified from an *Escherichia coli* strain that overexpresses the protein. The RNase Cocktail is free of DNase and nicking activities and is suitable for all appli-

cations where it is desirable to degrade RNA (for example, minipreps), says Ambion. Typically, 2.5 μ l RNase Cocktail is added to a 50 μ l plasmid miniprep before or concurrent with restriction enzyme digestion. RNase A cuts after C and U residues. Digestion of RNA with RNase A alone leaves fragments of RNA that are large enough to be visible on agarose gels and precipitate in ethanol. RNase T1 cuts after G residues. Consequently, the mixture of both enzymes results in a reduction in RNA fragment size over the use of RNase A alone. The Cocktail is supplied in 50 per cent glycerol, so that it remains liquid in the freezer and can be used immediately. It contains 1 mg ml⁻¹ RNase A and 20,000 units per ml RNase T1. Purified RNase T1 can also be purchased separately.

To overcome some of the problems associated with using cosmid or yeast artificial chromosome systems, a novel method for **cloning and packaging DNA fragments** using a bacteriophage P1 system has been developed by NEN Research Products and is available from Du Pont (*Reader Service No. 105*). The NEN-Phage P1 DNA Cloning and Packaging Systems allow the user to clone large genomic DNA fragments of between 85–100 kb in size with transformation efficiencies approaching those of cosmids, says the company. The P1 DNA Packaging System uses *Escherichia coli* strains and *in vitro* packaging extracts obtained from strains that are deficient in restriction and recombination abilities. These prevent the degradation and recombination of methylated genomic DNAs. The NEN-Phage P1 Cloning and Packaging Systems can either be purchased separately for £150 (UK) and £199 (UK), respectively, or as a £299 (UK) Complete DNA Cloning and Packaging System with sufficient reagents for five reactions.

Hans Landgraf has developed the Varius V **thermal cycler**, a multiuser system that features five heatblocks in one instrument, each of which can be independently microprocessor controlled (*Reader Service No. 106*). Two models are available: the Varius V₂₀ that holds 20 0.5-ml reaction tubes and the Varius V₄₅ that holds 45 tubes. The combination of light-weight parts (blocks of Sterling silver) with Peltier elements of high efficiency makes rapid temperature changes of 2–4 °C per second over a regulated temperature range of 1–99 °C ±0.5 °C



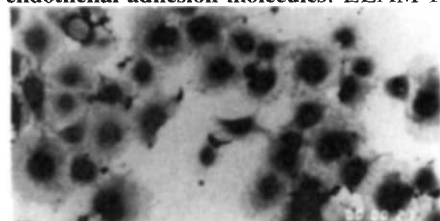
Varius V — five cyclers in one.

possible, says the company. The block uniformity is stated as $\pm 1^\circ\text{C}$. Other design features of Varius include an acoustic alarm system, storage of up to 30 programs, a parallel port for printing program parameters and a serial port for the attachment of an IBM PC or compatible computer.

The GeneAmp PCR System 9600 from Perkin-Elmer Cetus, which represents **second generation PCR technology**, is designed to improve speed, sensitivity, uniformity and convenience (*Reader Service No. 107*). The upgraded system integrates high-performance PCR protocols, thin-walled MicroAmp reaction tubes, a special sample block design and a 96-well, microplate-compatible format. According to Perkin-Elmer Cetus, the new system can perform the entire GeneAmp PCR process with an average three- to four-fold reduction in time and an amplification yield of 100,000-fold of the control target DNA without the use of oil overlays. The thin-walled MicroAmp reaction tubes are designed to maximize sample temperature uniformity by allowing the rapid heating and cooling of samples between 5 μl and 100 μl . The self-contained heating, cooling and control system regulates temperatures over a range of 4–99.9 $^\circ\text{C}$ with a resolution of 0.1 $^\circ\text{C}$. Up to 150 thermal profiles can be stored and existing PCR profiles from earlier models can be modified and transported into the new system.

Research tools: monoclonals

British Bio-technology has developed a new line of products designed to aid research into the function of various novel **endothelial adhesion molecules**: ELAM-1



Inhibition of U937 adhesion to COS cells expressing the ELAM gene by pre-incubation with anti-ELAM antibody.

(endothelial leukocyte adhesion molecule-1), ICAM-1 (intercellular adhesion molecule-1), VCAM (vascular cell adhesion molecule) and PECAM (platelet endothelial cell adhesion molecule) (*Reader Service No. 108*). The endothelial adhesion molecules are expressed on the internal surface of blood vessels (endothelial cells) and allow leukocytes, which are normally free in the blood circulation, to attach to the vessel walls. Under normal circumstances, most leukocytes remain in the circulation. But, in response to inflammation, leukocytes migrate through the blood vessel wall into the extravascular tissue to combat damage. Endothelial adhesion molecules are thought to be important in the localized extravasation of leukocytes and, as such, are likely to play a role in the inflammatory process and tumour metastasis. BBT's product line includes: monoclonal antibodies, which are suitable for immunocytochemistry and adhesion blockade; full length cDNAs — Designer Genes supplied in the expression vector pCDM8; and DNA probe cocktails intended for the detection of mRNA and DNA in techniques such as *in situ* hybridization and northern/Southern blotting.

The **epitope addition method of protein purification** uses BABCO's specific monoclonal antibody of known peptide specificity, 12CA5, to purify fusion proteins, including the epitope to which the antibody binds (*Reader Service No. 109*). The 12CA5 antibody was originally developed at the Research Institute of Scripps Clinic. The procedure is as follows: fusion proteins are expressed from a vector that encodes a specific nine-amino acid peptide epitope at the terminus of the protein of interest. The 12CA5 antibody directed against this epitope is then used to purify the fusion protein. According to BABCO, this approach has several advantages: the small sized epitope is unlikely to alter the reactivity of the cloned sequence; convenient vectors encoding this sequence have been generated; and the harsh conditions of extreme pH or chaotropic agents used in affinity chromatography are avoided.

T Cell Sciences has introduced the Diversi-T $\alpha\beta$ T-Cell Antigen Receptor (TCR) Screening Panel 1F, which detects α or β chain variable region epitopes (*Reader Service No. 110*). The panel, which consists of a collection of seven **FITC-conjugated monoclonal antibodies to human TCR variable region epitopes**, including V β 5, V β 6, V β 8, V β 12 and V α 2, can be used to determine the role of T cells in autoimmune diseases, cancer and viral infections. Each antibody is supplied in a ready-to-use form for direct flow cytometry analyses, and identifies an α or β variable region found on minor populations of normal peripheral blood



FITC-conjugated mAbs to human T-cell antigen receptor variable regions.

lymphocytes (1–5 per cent). As these populations may vary significantly from individual to individual, this panel can be used to investigate and define the human TCR repertoire in normal and diseased states.

Sera-Lab has introduced a **complement-fixing monoclonal antibody**, MAS 516, which is intended to be a useful research tool for cleaning up fibroblasts from human cell culture systems (*Reader Service No. 111*). Fibroblasts are often difficult to eliminate from epithelial cell primary cultures, such as thymus and breast, says Sera-Lab, and, as a result, epithelial cells can be overgrown by fibroblasts. MAS 516 reacts with the surface membrane molecules of human fibroblasts, causing lysis and, hence a reduction in the fibroblast population. The antibody will also bind to tissue macrophages and peripheral blood monocytes. Other stated research applications include immunocytochemistry on frozen sections, cytocentrifuge preparations and cell cultures (human fibroblast cell lines). MAS 516 is also suitable for ^{51}Cr complement-mediated cytotoxicity assays.

In kit form

Two new AvidChrom kits are available from BioProbe International for the **isolation of Fab and F(ab')₂ fragments** from antibodies (*Reader Service No. 112*). BioProbe claims that the AvidChrom Fab kit can produce Fab fragments from antibodies with yields of up to 80 per cent in as little as 90 minutes. The F(ab')₂ kit is designed to produce F(ab')₂ fragments with yields of up to 75 per cent in less than six hours; this is a considerable improvement over the use of immobilized pepsin



BioProbe's AvidChrom kit for high yields of F(ab')₂ fragments in less time.

systems, which produce yields of only 20–30 per cent in eight hours or more, says the company. The AvidChrom kits use a novel free-enzyme system, developed by BioProbe, which is designed to provide a more rapid and complete breakdown of antibody as compared to immobilized enzyme systems. The protocols are optimized for the removal of free enzyme, Fc fragments and undigested antibodies. Each AvidChrom kit contains sufficient reagents for ten digestions of 10 mg each.

The microplate immunoassay kit for human inhibin from the Brussels-based company, Medgenix Diagnostics, is designed to enable human inhibin concentrations in serum and other biological fluids to be measured in five hours (*Reader Service No. 113*). The two-site immunoenzymetric assay has a sensitivity of 0.1 units per ml, says the company. No interference was observed with TGF β , LH, FSH, activin and seminal inhibin-like peptide. The immunoassay kit may be useful in studies involving the physiology and physiopathology of human inhibin secretion.

Software: mini-profiles

GeneJockey, the latest software to emerge from the Biosoft stable, is a multi-

window software package for editing, manipulating and analysing nucleic acid and protein sequences on the Apple Macintosh (*Reader Service No. 114*). The £250 (UK), \$499 (US) GeneJockey program can rapidly scan nucleic acid sequences to select all possible pairs of oligonucleotides that are suitable for use as primers to direct efficient DNA amplification by the polymerase chain reaction. Searches can be made in GenBank, EMBL and other libraries, as well as GeneJockey's own files. In addition, nucleic acid and protein sequences can be imported from, or exported to, other programs. Restriction analysis can be performed, with recognition sequences given for over 400 restriction enzymes; more can be added. The program displays both sequence data and annotations in the same window, and allows up to 30 windows to be open at the same time so that sequences can be assembled from several sources.

ChemMod II is an interactive 3-D molecular modelling package from Fraser Williams that is designed for the Apple Macintosh II (*Reader Service No. 115*). The program runs under the UNIX operating system, which allows the user to handle large molecules without many of the constraints imposed when using MS-DOS, says the company. Files can be passed between the UNIX and Macintosh



ChemMod II — Mac-based 3-D molecular modelling software from Fraser Williams.

environments, and chemical structures can be integrated into Mac-based desktop publishing packages such as Aldus Pagemaker. ChemMod II provides comprehensive structure building facilities, including a diverse set of template libraries to assist the user in constructing energy-minimised molecules. Calculation facilities include full measurement of bond lengths, bond angles, torsion angles, di-hedral angles and distance, torsion angles versus potential energy plots and rapid Ramachandran mapping. Software for a single-user system starts at £1,500 (UK), which includes software updates and on-line user support.

Laboratory hardware

When used in combination with the Dura-Dry Microprocessor-Controlled (MP)

Condenser Module, the Dura-Stop MP Stopping Freeze Dryer from FTS Systems provides an automated lyophilization system (*Reader Service No. 116*). The user selects one of four modes of operation, enters the sample information and begins the run. Temperature and pressure are controlled and monitored by the Dura-Stop MP. FTS Systems offers accompanying Lyphware computer software, which provides additional program choices, increased control over run parameters and the option of data graphics.

Make microbiological media on an 'as needed' basis with the microprocessor-controlled microwave sterilization system called MSS-500, which is available from CEM (*Reader Service No. 117*). The MSS-500 can be used to prepare and sterilize media for up to 100 plates in ten minutes, says CEM. Dehydrated media and water are placed in high-pressure vessels, which contain magnetic stirrers to ensure thorough mixing. The system can accommodate up to 12 vessels with a total capacity of 1,200 ml.

Cecil Instruments has launched the Series 2000 range of UV/Visible spectrophotometers (*Reader Service No. 118*). The spectrophotometers are designed to provide a wavelength setting reproducibility of 0.1 nm, with straylight of 0.01 per cent, says the company. All functions, including peak seek and ratio of absorbances at two wavelengths, which are offered as standard, can be selected without the user having to resort to the instruction manual, claims Cecil. Remote operation is possible with the bi-directional RS232C port. The new design provides for an integral printer and incorporates a large cell compartment that can accommodate a range of accessories. □

These notes are compiled by Diane Gershon from information provided by the manufacturers. To obtain further details, use the reader service card bound inside the journal. Prices quoted are sometimes nominal, and apply only within the country indicated.

CORRECTION

The product review from Life Technologies (see *Nature* 348, 563, 6 December 1990) should have included the following disclaimer. The polymerase chain reaction (PCR) process is covered by US patents issued to Cetus Corporation. The customer is hereby notified that purchase of the SuperScript Preamplification System product from Life Technologies, Inc., does not convey to the customer any licence or authorization to practice PCR under any patents issued to Cetus. A licence to use the PCR process for certain research purposes accompanies the purchase and use of certain Perkin-Elmer Cetus products, or is available from Cetus.

ADVERTISEMENTS

For the Centennial of the Station marine d'Endoume (1889–1989) and from the related Symposium, the Centre d'Océanologie de Marseille has published "OCEANOLOGIE, actualité et prospective", M. Denis ed., 387 p. (ISBN 2-907752-00-6). Contributions in french and english from J.M. PERES, T.T. PACKARD *et al.*, J.F. MINSTER, L. LAUBIER, J.C. DUPLESSY, H. CHAMLEY, D. RAYNAUD, R. MARGALEF, J.C.J. NIHOUL, A. LAUREC, A.R. LONGHURST, R.C. DUGDALE & F.P. WILKERSON, A. MOREL, T. PLATT *et al.*, J.B. LEWIS.

Orders to AMADOU, C. O. M., Station marine d'Endoume, rue de la Batterie des Lions, F 13007 Marseille, France. FAX: (33) 91 04 16 35. Price 200 FF + 20 FF for mailing.

Reader Service No.7

METABOLITES

analysed rapidly, enzymatically, using micro-samples



with the

ANALOX GM7 MICRO-STAT

A wide range of biological fluids including whole blood are acceptable for direct injection. Turbidity no problem. Sample size 2.5 - 25ul. Results in 20 seconds.

glucose	lactate	urea
alcohol	pyruvate	creatinine
cholesterol	ammonia	3-OH butyrate
uric acid		acetoacetate

Analox Instruments Ltd

8 Goldhawk Industrial Estate
Brackenhurst Road Tel: 081-749 7635
Hammersmith, London W6 0BQ Fax: 081-740 6608

Reader Service No.6